From the Grid to the Smart Grid, Topologically

Pagani, Giuliano Andrea

From the Grid to the Smart Grid, Topologically

Giuliano Andrea Pagani

The work is supported by the University of Groningen under the Ubbo Emmius fellowship program and by IBM under the IBM Ph.D. Fellowship 2013-14.

Printed by *NetzoDruk* - www.netzodruk.nl - Groningen

rijksuniversiteit
groningen

From the Grid to the Smart Grid, Topologically

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus Prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 14 maart 2014
om 12.45 uur

door

Giuliano Andrea Pagani

geboren op 29 maart 1981
te Parma, Italië

Promotor:	Prof. dr. M. Aiello
Beoordelingscommissie:	Prof. dr. P. Pelacchi Prof. dr. F. Reed-Tsochas Prof. dr. M. van Steen

ISBN: 978-90-367-6813-9
ISBN-Electronic: 978-90-367-6812-2

To my family

Contents

Acknowledgments

Once a friend asked me what is like to do a Ph.D. and my answer was something like: it is like you are driving a ship in the ocean at night with a broken compass in a misty weather, but then you keep going and after a while you look at the sky and see the stars that can guide you to your destination ashore.

The Polaris (i.e., North Star) of my journey is represented by my promoter Marco Aiello, to whom my deepest and greatest gratitude goes. He has been the inspiration of this project with his highly creative mind, he helped me a lot in process of make this Ph.D. journey in Groningen to happen. The Ph.D. journey is also full of emotional challenges that range between happiness, fear, delusion, and anger: a full swing of emotions. Marco was always there to joy with me in the good moments and ready to define new goals and challenges, as well as encouraging me in the bad times with words full of positivity and hope and good pieces of advice. Thank you very much for all that, not least the pleasure of talks in Italian sometimes.

A special thank goes to the stars that have guided me in the last the final effort of writing the logbook (i.e., this thesis work). Thanks Paolo Pelacchi, Felix Reed-Tsochas, and Maarten van Steen for the time and effort of reading and evaluating my thesis work.

Other shining stars during the navigation process that contributed to my scientific enrichment are represented by Ettore Bompard, Martí Rosas-Casals, and Antonio Scala. Thank you all for the fruitful exchange of ideas in the area of complex networks and power grid infrastructures. Other fruitful discussions regarding the area of energy market and implementation platforms took place with Nicola Capodieci and Giacomo Cabri, with whom I enjoyed the pleasure of successfully collaborating in various scientific publications. I really enjoyed the joint work with my colleagues Viktoriya Degeler, Ilche Georgievski, Alexander Lazovik, and Tuan Anh Nguyen; it was really fun to work together on a bigger project and mess up with the equipment in our office space.

I also thank the constellation of friends and then co-authors of the Santa Fe Institute Complex Systems Summer School 2013: Holly Arnold, David Masad, Johannes Schmidt, and Elena Stepanova. Our interesting talks and interactions, both on site in Santa Fe and via Skype, are really important to develop novel challenging ideas in a new territory of network science; thanks for being part of the NetAttack group. Another Santa Fe star is Kerstin Honey with whom I started with lots of enthusiasm the exploration of the *terra incognita* of energy and biomimicry.

The sky is full of stars as well as the many colleagues with whom I shared the Ph.D. journey, not only in the Bernoulliborg, but also in group meetings and travels. Special thanks to my office mates for long time Pavel Bulanov and Ando Emerencia. Other present and former colleagues of the group deserve a special mention: Saleem Anwar, George Azzopardi, Frank Blaauw, Doina Bucur, Kerstin Bunte, Mahir Can Doğanay, Elie El-Khoury, Ilche Georgievsky, Ioannis Giotis, Heerko Groefsema, Eirini Kaldeli, Alexander Lazovik, Tuan Ahn Nguyen, Faris Nizamic, Fatima al-Saif, and Ehsan Ullah Warriach.

Of course, no navigation is possible without the supporting staff that provides you with the appropriate equipment, advice and material. A great thank you or 'dank jullie wel' to the secretary team: Esmee Elshof, Desiree Hansen, and Ineke Schelhaas. Thanks for managing a lot of administrative burden and paperwork for me, not less important thanks for the nice opportunity to practice my Dutch.

A special thanks to those people that provided me with the maps to read the stars. This research would not have been possible without the data, help, suggestions, and indication of Hemmo Hulzebos and Berto Jansen. Peter Kamphuis was another key figure, involving me in several projects at his institutions and providing me with the right contacts in the utility world. Another special thank goes to Anne Beaulieu for proposing me collaborations in interesting projects related to the energy topic.

Although the Ph.D. journey is a lonely sailing experience sometimes you cross the route of other sailors. Between the many friends busy with their journeys a special mention goes to the small piece of Italy I hung-out sometimes: Sissi Balbo, you are a special friend and we have to write an article about networks in your field of study together, that was our deal; Elisabetta De Cao and Nicola Barban, thanks for the nice all-around conversations. A special though goes to the friends at UMCG with whom I shared many week-end evenings talking about research, sharing Ph.D. life and beyond, thank you Esmee Joosten, Piotr Kowalski, Malgosia Krajewska, Monika Maleszewska and Milind Pore. Another thank goes to the APSAZ volleyball club for the friendly atmosphere and for making me re-discover my volleyball skills. Special mention to Chakir and Sabrina Bouziane, Mark Hilbrands, and Paul van Snick for the volley performances and the subsequent nightlife. Another mention to people that in different times were important in my 4 years in Groningen: thanks to

Eva Koning among the many things the lessons about horses, animal guidance and leadership; thanks to Steffannie Wiekens for the humor and the nice talks and chats; thanks to Estelle Meima for being my tennis partner and for the nice conversations afterwards.

A special thank to my fellows sailors on board (i.e., paranimphs) Ando Emerencia and Matt (Mario) Heberling. Ando has been a good office mate with whom I shared the whole trajectory of the Ph.D. experience; a special thanks for cross-checking my samenvatting. Thanks to Matt for the nice runs together and for the deep (and not so deep) conversations when sharing my house together. I also really appreciate the proof read of some key chapters of the this thesis and the adventures during 'tiger nights', thanks Mario.

Of course no sailing trip is possible without a ship owner that takes care and provides you with all the material and moral support to undertake the journey. The ship owner was always in contact with me thank to an on-board Marconi's telegraph (evolved version). Thank you to my family for the constant support in these four years. I am forever grateful for those close to my heart and although this vault of the sky is now deprived of the shine from their star, their influence on my life is special and enduring because it still feeds my heart with courage and warmth. Special thoughts to my grand mothers Carla and Gabriella for their phone calls and interests in my life and for making me feel in the warm settings of my home life. A special thank to my father Gaetano for the profound thoughts and pieces of advice in several evenings; and for the new relationship between us (Altima trips, etc.). A very big thank you to my mother Margherita, you were always close to me in any sense: several relocations, financial support, and in the days of my worst mood to cheer me up. Thanks for the help in many occasion and the nice and comforting words and pieces of advice.

And the last thought goes to the person ashore who is shining like the Morning Star at the end of the dark night journey, and who is looking forward to start new journeys with me. Thank you Feikje for being with me, for your nice and sweet words all the time and for the help with 'Nederlands vertaal'. I really look forward to start new journeys and adventures together.

G.A. Pagani
Groningen
February 6, 2014

Chapter 1

Introduction

The human body of a man 25 years old at 180 cm and 75 kg, requires just about 2.128 kWh of energy (1830 kcal) a day for his basal metabolic rate [85]. All other energy consumption that each of us requires is to make our life more productive, more enjoyable, and more safe-in essence, more rich. Energy is the enabler of modern societies and cities. If one looks around, everything is powered, moved or lit by some form of energy. The availability of energy is taken for granted in the modern society so much, that we perceive its real importance only when outages or blackout strike. In fact, the effects of prolonged blackouts have threatening consequences both for population security and safety, as well as for economic production [7, 11]. In developed countries, energy has become such an affordable and easily accessible commodity that we are reminded of its presence only when the car needs to be filled or when we receive the bill of the energy retail company.

The scope of the present work is on the electric grid and system, therefore, when we refer to energy we typically mean it in its electric form. Actually, electricity is not the only form of energy that was mainly used. This was even more true in the period before the advent of the Information and Communication Technology (ICT) and the digital life. Worldwide, in 2011, the total consumption of energy is dominated by oil and then electricity with 40.8% and 17.7%, respectively; followed by natural gas (15.5%) [90]. If we look back to 1973, for example, we note that in 2011 we have an almost doubling in the total consumption (4674 million tons of oil equivalent vs. 8918 million tons of oil equivalent); and looking at the share composition in 1973, oil was even more important with 48.1% followed by natural gas and coal with 14% and 13.7% of total consumption, respectively; whereas electricity had just a 9.4% share [90].

Electricity applications were initially not completely understood and the effects of electricity were not considered particularly useful outside of the physics lab. The first application dates back to the incandescent bulb. However, here electricity is a "carrier" of energy that is transformed in other forms of energy, namely, thermal energy, and light. The first application of electricity, per se, has to deal with the telegraph. Again, electricity is used as a carrier this time for the electric signals encoding of information. Electricity was then used mainly as a carrier of energy from a source point (i.e., power plant) to the end user. Usually at the source, various other forms of energy are converted into electricity (e.g., the conversion route may entail be combustible fuel, to thermal energy, to motion, to electricity). Once in form of electricity, energy can be transmitted from the source to the destination through a capillary infrastructure: the power grid.

The power grid is considered one of the greatest achievements of the 19th and 20th centuries. Just consider the challenge of bringing the power through a whole continent, such as Europe or America. A typical simplified scheme of the power grid is represented in Figure 1.1. One sees an organization of a transmission (extra high and high voltage) and a distribution grid (medium and low voltage). Energy is mainly produced in large power plant facilities by a few authorized actors, usually far from big population or industrial conglomerates where the energy is needed. As the plants are the components at maximum power in the system, they are connected to the high voltage level; while end users consume mostly at the medium and low voltage levels. The technical reason for this voltage stepping is to prevent the dissipation and losses for the transmission of the high amount of power with small currents, and safety in the surrounding of the end user. In fact, losses are proportional to the square of the current. Therefore, on the one hand, very high voltages are required to transmit power while having small currents. Whereas, on the other hand, high voltages are deadly for humans, thus, requiring low voltages in residential premises. The structure is therefore highly hierarchical. This is also reflected in its commercial organization, where, at the beginning of the electricity era, there was a monopoly for regulatory, technical, and economic reasons. Samuell Insull (1859-1938), initially an assistant to Edison and then an energy entrepreneur himself, stated a famous quote that summarizes the approach to the electricity sector in the early days [131]:

> *"Every home, every factory, and every transportation line will obtain its energy from one common source, for the simple reason that that will be the cheapest way to produce and distribute it."*

The concept is perfectly correct in its historical context and in an electrical system where energy sources are cheap, cost of transmission are negligible and environmental externalities for energy production are not taken into account. However, when these factors start to take their matter, the electricity sector might begin to consider different paradigms and technologies for production and distribution.

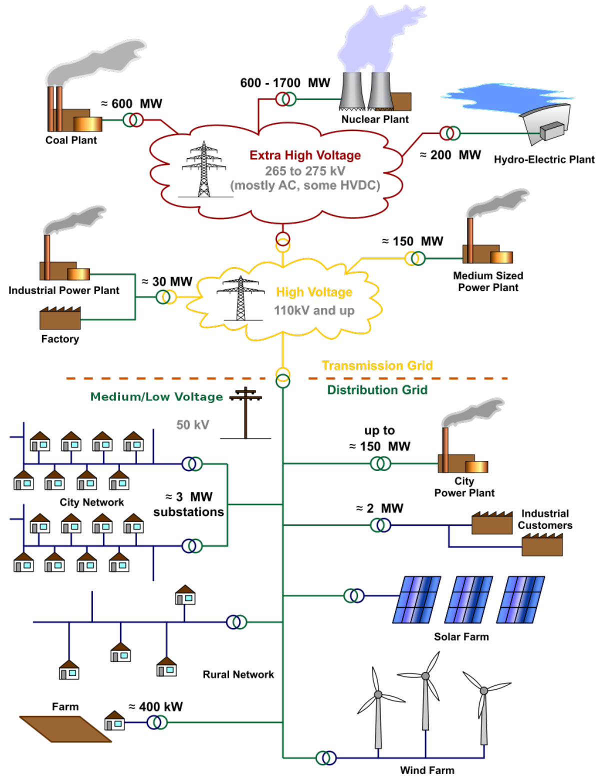


Figure 1.1: The physical organization of the power grid. *(Source: adapted from Wikipedia)*

1.1 The New Grid

During its evolution, the electrical system has become more and more pervasive, and ICT and telecontrol have been applied to monitor, control, and actuate; especially in the high voltage part of the system. The medium and low voltage end of the grid have received less attention due to their traditional role as a passive component of the grid. With the current technology and operation of the grid in the western world, the power supply is highly reliable; for instance in the Netherlands in 2012, the electrical system was 99,99486% of the time available with just an average of 27 minutes downtime per year per customer [149]. The Dutch system has one the highest availability in Europe. Even though the outlook seems good and with no issues, the future is not trouble-free.

In a 2003 meeting involving American utilities companies and U.S. governmental institutions [207], the concluding outlook on the status of the grid was not reassuring: the grid was found outdated with limited possibilities of expansion in the transmission lines close to urban centers, inefficient, and unable to meet the energy requirements of the information-oriented society and economy. Low investments were realized especially at the level of distribution, and many projects with the aim of increasing the capacity were canceled. However, some encouraging points emerged from the meeting. Examples include the focus on new technologies including smart power systems and clean energy systems. In addition, investments on the order of billions of dollars to replace existing equipment and a plan for new ones were considered. Another aspect mentioned in the meeting called for clearer rules for regulation. In essence, the report of the meeting laid the foundation for a smart grid.

1.1.1 Motivations for the future grid

The motivations that drive the requirement to update and modify the traditional system lie in several points. The first point of concern is the reduction of greenhouse gas emissions: the system has to accommodate large quantities of sustainable and clean sources of energy. This aspect poses a challenge to the traditional energy system. Actually, renewable energy poses threats to the electrical system due to their difficulty in output predictability. In fact, the electric system was designed with the idea in mind to follow the demand by using resources with a predictable output, such as fossil fuels, and also having the possibility to dispatch a plant. Dispatching is the operation of planning the use of a power plant in order to meet an expected future demand. Different plants are dispatched at different moments in order to satisfy demand at any time and given the different inertia of the power plant with different technologies. Plants with high inertia satisfy the base load (e.g., nuclear and coal), while plants with smaller inertia satisfy the medium and peak load (e.g., gas and oil). Renewable energies pose a challenge since they highly depend on natural variability. Therefore their prediction is not accurate (exceptions are hydro, biomass and tidal powered plants) and it is clearly impossible to drive nature to follow human demand. One can see the difficulty in having a system fully based on renewables with the classic idea of demand following [72]. The system needs, therefore, to be more flexible to accommodate the energy demand based on the availability of the renewable resources in a point in time. A second motivation for the future grid is the technological and economic availability of small generators producing energy from renewable sources such as photovoltaic (PV) panels and small wind turbines. These technologies are nowadays often subsidized by governments with the aim of reducing greenhouse gas emission and to achieve higher penetration of renewable energy. A third motivation lies in the transformation that the energy market is undergoing. From the initial monopoly, the situation has changed in many countries and now more and more companies are present in the energy business from energy production, to energy transmission and distribution, to retail and services provided to the end user. For instance, this is the case of the Netherlands. To enable and accelerate this process, governments in the western world have promoted policies to open the electricity business and facilitate competition with the final aim to both modernize the energy sector and provide a more convenient service for the end user. Such tendency of the last decades is termed *unbundling*. In essence, unbundling is the process of dismantling monopolistic and oligarchic energy systems, by allowing a greater number of parties to operate in a certain role of the energy sector and market [51, 93]. Considering the vision of combining the previous two aspects, one notes that it is possible for everyone to become a small producer and energy provider (of small-scale). Such a small-scale approach is considered beneficial to the electricity system in many ways: from reduced losses, since source and load are closer; to system modularity; to smaller investments compared to large-scale energy solutions [123]. Another motivation towards a modernization of the grid lies in the necessity of using modern ICT technologies for the operations of the energy sector. ICT and automation can ease many large scale operations that were (and still are) done manually such as meter reading, user connection and disconnection, and network switch operations. In addition, by using ICT in the energy domain provides new possibilities and scenarios for the end user, one can raise energy awareness and provide automation for energy efficiency.

1.1.2 What is the smart grid?

Smart grid is a recent term that does not have a unique single meaning. This uncertainty is apparent in the statement by the The Global Community for Sustainable Energy Professionals - Leonardo ENERGY[1] *"A smart grid is neither a clearly defined single concept, nor a single technology. Rather, it is like a basket containing various combinations of balls. The context and the interpretation depend upon the user."* More specifically, the Department of Engineering and Public Policy of Carnegie Mellon University describes that the characterization of the smart grid is non-unique and various aspects are involved in the smart grid concept [139]. According to the authors, there is not a single meaning for the new energy landscape concept, but rather every actor involved in the grid has different views of the smart grid according to his related business; and the various stakeholders of the smart grid landscape can benefit in different ways from it:

[1] http://www.leonardo-energy.org/what-definition-smart-grid

- The customer can benefit from real-time tariffs that reflect the price of electricity, react with his loads to the tariffs and receive information directly from the meter about his consumption and costs.
- The electricity distribution companies can have a higher level of automation to manage ordinary operations, critical situations, and a more selective way of shedding load (e.g., based on the importance of the service provided: hospitals or police stations might be the last to be removed from the network due to their social and safety importance). For distribution companies, the smart grid also enables scenarios to manage easier integration of distributed energy generation facilities in the low level grid (e.g., photovoltaic installations and wind farms) or enable islanding of subsets of the grid during emergencies.
- The generating and transmission companies can benefit from a more "computerized" grid with more information and data about critical grid measures (e.g., network's phase voltage). Having more information enables more automatic and distributed decision-making even far from the control center, thus optimizing the grid operations.

Now a natural question concerns the actors of this system. The scenario is more complex when a monopoly is not in place due to higher competition, efficiency requirements, and a different scale of the operations involved. The view of National Institute of Standards and Technology (NIST) for smart grid described in the NIST Smart Grid Framework [156] is depicted in Figure 1.2, where the many actors and their domains in the smart grid landscape are represented. It is thus clear from the picture that many different actors are present and that many different technologies need to interact at different levels to properly enable the synergy between the many actors. In the figure, one notes that what is really new is the information flow between the various parties involved. Traditional flows of information that involve the generation and transmission grid with the control operation center and the market are still present. These flows of information are essential for the monitoring, control and operation of the grid; they have been a key component of the grid for a long time. We now see that the distribution grid is also involved with information exchange with the market, the operations, and the customer. The customer becomes an active player and his/her appliances and equipment will be monitored by the utility service provider for commercial purposes and at an aggregate level by the distributor for metering and technical purposes (e.g., power quality). We also see the interaction directly with the market that enables the user to become an active participant in selling and buying his own energy. The techniques used to interact with the end user to modify his load are broadly named demand-side integration. Demand-side integration incorporates various other terms that stress

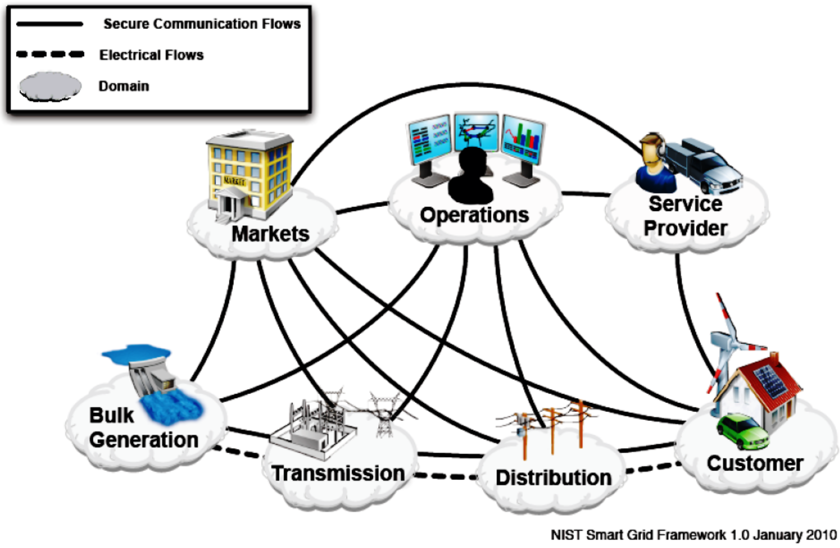


Figure 1.2: NIST smart grid framework.

some peculiar aspects, such as demand-side management, where utilities companies perform activities to influence customer energy behavior (e.g., peak/off-peak tariffs); demand-response, where the demand is shaped to respond to supply conditions (e.g., real-time pricing); demand-side participation where users participate in three aspects, providing a more economic, secure, and environmental-friendly electric system [66]. Depending on the type of customer, utilities implement different methods to control or influence customer's load, such as direct load control where the utility controls the customer loads (e.g., air conditioners, water heaters), or agreeing on a curtailable service during contingencies that give the customer a discount (usually applied to big industrial or commercial customers), or demand-side bidding where customers offer curtailment with bids based on the wholesale energy market. Other operational implementations consider the situation of emergency curtailment and ancillary services for market; in this last one, the customer bids to stay on a standby situation and curtail if needed. Looking at the energy flow, we see that it still follows its original route from the generators to the customer. However, in the last segment of the grid (i.e., distribution), the customer has become a producer (cf. PV panels and wind turbine on top of the customer house in Figure 1.2) that feeds the energy surplus back into the grid. Actually, a new name has been coined to indicate the new figure of end user that has both the roles of producer and consumer of energy: the prosumer.

In summary, smart grid is not a single technology, but rather an approach involving multiple technologies to improve and change the way the actual power grid is used and managed. The most important innovative aspect is to have, in parallel with the energy flow, an information flow that enables more advanced functionalities both for the grid operators and for the end users. A synthesis of the smart grid approach can be seen in the reports issued by the U.S. Department of Energy [145] and [146] with the main underlying idea is that the grid should be able to:

- self-diagnose and take appropriate recovery actions in case of faults,
- become more resistant to unintentional and malicious attacks,
- satisfy the users with an improved power quality meeting their needs and expectations,
- be ready to integrate different sources of power generation,
- give the end users the possibility to interact and respond to real-time electricity price signals in a distributed energy market.

1.2 Thesis Scope and Organization

In the vast smart grid domain, the research challenge that we tackle in the present work concerns the changes necessary to fully enable the smart grid vision. The first research question deals with the infrastructure. Which development must the physical power grid infrastructure undergo to support a future scenario where users can locally produce their own energy and are free to participate as prosumers in an unbundled energy market? The second topic of interest deals with the ICT applications for the smart grid and leads to the following research questions: which architecture and features of future applications will enable the smart grid? What are the application scenarios available for the user in a fully unbundled energy market?

To answer the first question, we focus on the part of the power grid infrastructure that is close to the end user: the distribution grid. We resort to complex network analysis to analyze the current state of the grid infrastructure and consider possible evolution in the future. Complex networks are a key framework to better understand those systems that are composed by many interacting parts in a network fashion in order to grasp an overall behavior of the system. The study of complex networks is part of complex systems theory, which is the multidisciplinary scientific branch that studies systems with many interacting parts (or agents) in which the overall emerging behavior of the system cannot be deduced by the observation of the single parts. Other essential features of complex systems are the absence of a centralized controller and the evolution of the system over time [136]. The choice of complex systems approach is twofold: first, since the electrical system might evolve towards a complex system; second, at the present stage, given the major uncertainty surrounding the smart grid topic, the planning of the new grid is not to be a fully detailed one, but more of a high level scenario where coarse grain network models are useful instead of circuits at very precise level of details. Having in mind these characteristics of complex systems, we envision that with more autonomy in production and consumption, the system might reduce the centralization control and go towards an electrical system that is locally self-organized where users interact in selling and buying energy. Self-organization and lack of a central controller together with interaction between multiple entities are some of the fundamental principles of complex systems theory. The analysis of networked systems following the paradigm of complex network analysis has provided interesting results in several fields (e.g., computer science, health, biology). We consider complex networks as a useful tool for our goal of realizing a high level decision support system to study and evolve the smart grid which is a networked system by definition. In particular, having a weighted representation of the network, where the weights represent physical properties of the grid, is helpful in approaching the real infrastructure without all the very low levels details. Network science has its basis in graph theory and in the mathematical branch that study topology. When referring to networks, topology is defined as the particular organization (i.e., pattern) of the connections (i.e., links) between the nodes that compose the network.

To answer the questions more focused on the ICT application side, we focus on the software approach that is capable of sustaining a set of systems and applications that are diverse, highly distributed, and with constraints for security and timing. These characteristics are satisfied by service-oriented architectures (SOAs). Going more into the application domain of the smart grid, we aim at realizing and testing, in conditions and environments as close as possible to the real ones, the applications that will likely to be in place in the future smart grid. The success of the smart grid will depend on the applications that the end user will be provided to manage his energy production and consumption. Building and testing these applications in a living lab environment is the first step to their engineering to a higher scale.

These two topics are closely related: the smart grid is about an improvement of the infrastructure to make it more efficient, safe, and a rich of source of information, on the one hand, and enable new applications to promote flexibility in the energy system, on the other hand. The infrastructure and ICT application domains will have to interact. Figure 1.3 shows this overall picture. The grid and its components such as transformers, substations, small-scale and traditional generators, and lines will have, especially at distribution level, more ICT-based sensing and automation capabilities compared to the current situation. Each physical piece of equipment will also be characterized by a service point that it provides to other entities in the SOA for the smart grid. In the SOA, information layer logical entities provide and consume information in a standardized (e.g., eXtensible Markup Language-based) fashion. Logical entities can be supervisory control and data acquisition (SCADA) systems and energy management systems (EMS) belonging to utilities or to the end user household. Also the smart meters or building energy management systems, and the agent technology that will reside in them, will be a part of the SOA infrastructure and therefore the gateway to access and provide external services. Using such an approach, the electricity sector will have a substantial change in its operations being more ICT and communication intensive and, at the same time, more efficient and agile (e.g., automated metering reading, load shaping through dynamic pricing, switch automation). In addition, the smart grid has the objective of raising user awareness and interest concerning energy. With the smart grid, the energy sector is not the utility sector anymore where user pays the bill with the same spirit and will of paying income taxes, but a vibrant sector where new services and opportunities are ongoing towards a more sustainable production and use of energy.

The title of the present thesis mentions the term topologically. In fact, we want to stress the focus on the new patterns of connections and shapes of the new grid. We examine in detail how the topological infrastructure that characterizes the grid will have to change its connectivity in the transition to the smart grid. In addition, we will look also at the connections between new classes of objects that need to interact, therefore being connected, from an application point of view. These new objects that interact digitally are smart meters, SCADA systems, EMS, local energy markets, distributed power grid sensing equipment, devices and automation actuators inside the house. The interaction between all of them will build a new panorama for the energy sector.

The novelty of the thesis resides in the several aspects. First of all, in studying a new topic such as the smart grid, then in focusing on the distribution network that has not received much attention in the literature. In addition, the approach used is based on complex network analysis and our contribution proposes a new framework model of evolution towards the smart grid. Our study takes into account also the economic factors of the evolution of the grid and the infrastructural elements that influence the cost of electricity distribution. Our findings lay the foundation for a decision support system that helps the utilities and governments in considering evolution scenarios of the distribution infrastructure towards a local energy generation and exchange. On the software and application side of the smart grid, we investigate novel scenarios and applications available in the future electrical system; our main contribution is to have these application tested in real environments with

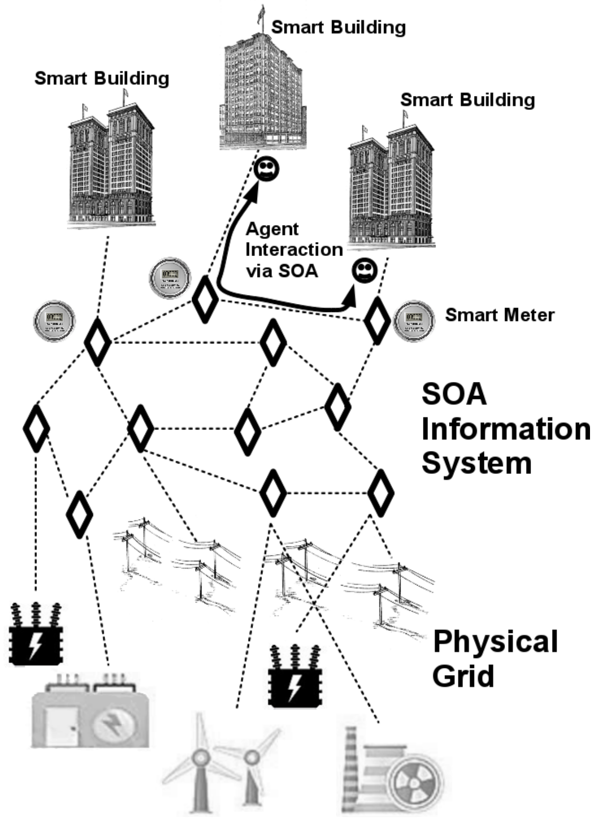


Figure 1.3: Smart grid physical and information infrastructure.

realistic data in order to prove the feasibility of the smart grid vision and concept.

The thesis, therefore, reflects the two main topics mentioned above: infrastructure analysis and design, and ICT applications tailored to the smart grid.

The thesis entails three focus areas: visions of the future energy system, complex network approach to the power grid, and software approach and applications for the smart grid. Chapter 2 provides the necessary background and state of the art. First, we focus on the vision and scenarios characterizing the new energy world enabled by the smart grid. Second, we review the most relevant work in the field of complex network analysis applied to the power grid considering different geographies and the focus of the different studies. We also study the main techniques of design and improvement of a grid infrastructure. In particular, we look at several aspects: the complex network analysis work relevant to the power grid design, the evolution of complex networks where links are missing or need to be added, and the design approach used in the electrical engineering field. Third, we look at the literature concerning the software approach for the smart grid services and their applications.

In Chapter 3, we analyze several samples of the Dutch medium and low voltage grid by using a complex network analysis approach. Since, in our vision of the smart grid, the end users are both producers and consumers, we focus on this layer of the grid close to the end user. We expect the medium and low voltage network to be crucial and invigorate lots of innovation, also in the topology of the infrastructure. The study of this layer of the power grid has not received particular attention before, especially using complex network analysis techniques. Our study takes into account not only purely topological features, but also physical properties of the lines. In addition to the topological analysis, we look at economic measures that influence the price of electricity distribution that are related to the topological aspects of the grid and apply them to the Dutch case.

Chapter 4 goes further in the complex network approach applied to the power distribution grid. We explore the evolution of the infrastructure towards the smart grid. First, we consider the development of a distribution grid infrastructure from scratch. Therefore, we look at the most famous network models from the literature and evaluate them against a set of topological metrics that are relevant in the smart grid context. We also present an evolution of the current Dutch grid samples where we propose an increase in connectivity of the network compared to the current one. We explore several strategies of adding new lines between existing nodes and we evaluate the evolved networks against the set of metrics. In both studies, we emphasize the economic side in terms of the cost to build such networks (or their evolution from the current infrastructures) and the benefits that the change in topology reflects on the electricity distribution costs.

Chapter 5 focuses on ICT applications. We look at the service-oriented approaches to the smart grid and point out the aspects that make it suitable for this kind of application. We then dig into the realization of relevant foreseen services that the smart grid will need to have. We provide an ideal solution for the future and we also show how a feasible implementation, close to the ideal, can be realized nowadays with current technologies and Internet services. An application of these real services for the schedule of appliance operations driven by energy dynamic pricing in an office environment is presented. The last application provides the future energy market. We present an agent-based platform to perform energy trading between consumers, prosumers, and traditional energy companies. We show how a bidding strategy for an agent can be optimized and how software agents can be coupled with real metering devices such as smart meters.

Chapter 6 concludes this thesis providing some discussion on the topics presented and a direction for the future work.

The thesis is concluded with two appendix chapters to provide essential background. Appendix A provides an overview of the most important graph theory and complex networks concepts. Appendix B provides a description of the most known network models from the complex networks literature that are used in Chapter 4.

1.3 Publications

Part of the work presented in this thesis has been published in or, at the time of writing, is under consideration for publication by several journals and conferences. In Table 1.1, we provide an overview of the papers on which this thesis is based and the chapters they are mostly relevant for. Additionally, the papers [157, 35, 34, 162] have contributed to shape the smart grid vision reported here, but are not directly related to any chapter. We stress that the contributions are to be considered a joint effort with the respective co-authors.

Chapter	Venue	Citation
2	Physica A: Statistical Mechanics and its Applications	[167]
	Service Oriented Computing and Applications	[161]
3	IEEE Transactions on Smart Grid	[159]
4	IEEE PES ISGT 2013	[166]
	ISCIS 2012	[163]
	Physica A: Statistical Mechanics and its Applications	[168]
	Submitted to journal	Tech. Rep. [164]
5	IEEE Transactions on Smart Grid	[76]
	ICAC 2011 Workshop Session	[36]
	ICSOC 2010 Workshop Session	[158]
	Submitted to journal	Tech. Rep. [165]

Table 1.1: Overview of publications and manuscripts under review related to the topics presented in the chapter of the thesis.

Chapter 2

Related Work

The smart grid is a vision and a trend for the electricity system as described in Chapter 1. It is also a new topic of research that in these past few years has received more and more attention. The reason is the ingenuity required to provide more flexibility to the electrical system to accommodate more renewable sources, to design new solutions to include electric vehicles, to consider users that are no longer passive and just consumers, but that are able to generate their own energy requirements and feed the surplus to the grid. In the past few years, the academic research devoted to the smart grid has increased, and a proof is the establishment of scientific journals on the topic and international conferences and workshop focused to this theme.

Here we provide an overview on the smart grid with special focus on the main aspects of this thesis: the vision of the future energy interactions; the study of the topology of the power grid and its evolution following the approach of complex network analysis (CNA); the status concerning applications involving Information and Communication Technology (ICT) technologies for the smart grid.

2.1 The Visions of New Grids

In the last decades, the view of considering the electrical system and the power grid managed and administered by a single company or institution has changed and questions to this model have appeared. More and more views of involvement of users in the energy picture are appearing. The idea of a new energy landscape is highlighted in the book “Power to the people” [208]. In the book, among the various solutions to achieve a more sustainable energy future and a viable way for our planet to avoid the effects of global warming, is a new paradigm where local (i.e., SM all-scale) production of energy and ICT are the cornerstone of the Internet of energy that is to come. According to the author it is not only technology driving this change, but its combination with rising environmental issues and the changes in the political/economic landscape fostering liberalization and competitiveness. The vision of an energy path that can can avoid big power plants and huge investments is also promoted by the environmentalist Amory Lovins in his book “Small is Profitable:

The Hidden Economic Benefits of Making Electrical Resources the Right Size" [123]. In the book, both the economic and the environmental factors are taken into account and the viability of solutions based on distributed generation powered by renewable sources is considered a better alternative than traditional ones characterized by pollution and green house gases emissions. Small-scale modular solutions are better than the traditional oversize power plants from a financial point of view given the delays and costs that such large projects incur. In addition, with modular distributed generation plants 207 benefits can be achieved according to Lovins. The same attitude towards the benefits from economic, environmental and reliability perspective in adopting micro-power generation (i.e., plants smaller than 10 MW) is shared by Dunn [62]. In another book titled "Hot, Flat, and Crowded: Why We Need a Green Revolution - And How It Can Renew America", Thomas Friedman sees the Internet of energy as the revolution that can help in reaching a sustainable future made of green energy together with energy efficiency and energy conservation [73]. As in the Internet, users create content and organize in communities, in the *Energy Internet* users can actively participate in production of energy and organize in virtual power plants or energy cooperatives. The benefits of an approach to the electrical system based on micro-grids rather than traditional macro-grids is provided by Marnay and Venkataramanan [129]. The authors emphasize three benefits achievable with micro-grids. First, micro-combined heat and power (micro-CHP) units have the ability to produce heat and electricity close to the location where it is needed the most. Second, power quality and reliability requirements differ among the pieces of equipment, performing distinct functions for the end user. Therefore, the authors speculate on the possibility of realizing different grids to be used to power equipment with different levels of quality and reliability. This idea is based on the criticality for society of certain equipment and thus the possibility (or not) to tolerate more instabilities. The vision of an electrical system more similar to the Energy Internet is proposed in [206], where the authors propose a system that in order to be self-healing and more efficient should be characterized by agents for every component or entity (e.g., meters, transformers, generation plants, customers). The agents can reconfigure and react to the changing conditions in the system. The authors explain three main differences between the data transmission and the electrical transmission: first, electricity is usually centrally generated and only one or few way of transmission are available, whereas data are generated and routed everywhere in the network without a priori fixed paths; second, it is difficult and inefficient to store energy in considerable quantities while data are easily stored and cached in the Internet; third, the whole Internet is based on the best-effort concept and quality of service is the exception, whereas the security of supply of the customer is the primary rule of the power system. The authors propose as a solution for a self-healing and a more flexible network, to have agents sharing information on the status of each component, precise customer load forecasting, and virtual electricity storage. This last concept is achieved through a demand-response approach intervening on customers' loads. The idea of a more decentralized grid in the generation and in the control of the production and the assets is expressed in [19]. The authors analyze peer-to-peer topologies adequate to the distributed infrastructure where agents (i.e., peers in the network) can control equipment and cooperatively decide on the best action to take when a problem is flooded to the agents. The agents also act on behalf of users in an online power market and can shed loads on behalf of the users when the grid is in critical conditions. We see similarities between this approach of considering the new interactions between agents on the grid and the peer-to-peer concept of the smart grid that we have described in Section 1. Our approach is more focused on the high level interaction between prosumers, while other works consider the peer-to-peer paradigm also for communication and control purposes [19]. The vision for a future energy scenario where energy services are the core of the smart grid is proposed in a vision paper about 10 years old by Gellings [75]. The author envisions a convergence of electricity, communication, and energy services in a device defined as consumer portal. This portal is the fusion of a smart meter, an Internet access router (the author considers a power line communication access to the Internet), a Web portal and a communication gateway to the intelligent appliances at home. Through the portal the energy companies (i) can promote energy-related services, (ii) can provide the user with tariffs, manage the status (i.e., metering) of other utilities services (e.g., water, gas), (iii) can monitor the appliances to inspect for failures, (iv) can monitor the status of distributed energy resources and energy storage devices. It is surprising how this vision that is on the scene for a decade it is still very up-to-date and basically today's ideas concerning the home energy management systems are unchanged and still far to a massive deployment at customer side. The idea of a home and building automation platform for energy purposes is the focus of Moneta *et al.* [137]. The vision follows those described above, where an energy market with multiple suppliers and freedom of choice for the customer are in place. Utilities offer additional services to the customer such as remote management of the distributed energy resources, the shedding of specific loads, where the customer can respond to changing tariffs or emergency conditions. The authors have developed a software platform that enables these functionalities and they have tested it in a realistic home environment.

In summary, the visions concerning the electricity system of the future are in line with our own view of the smart grid, as described in Section 1. The future of electricity systems will be much more focused on small-scale generation and distribution where users will be active participants in generation and balance of the system. ICT technology will make this change possible through smart meters and home energy management systems that are able to interact on a free energy market on behalf of the user. Users will be producers and likely to be energy feeders in local micro-grids. The energy utilities will be no more utilities, but will evolve towards energy service companies offering more energy-related added value services (e.g., appliance monitoring, energy efficiency). The grid, especially at medium and low voltage level will be more remotely managed and controlled with software agents that control power assets and interact together to decide the best and safest configuration for the equipment.

2.2 Complex Network Analysis and the Power Grid

The works that we examine consider the power grid networks as graphs following the mathematical meaning of the term. The main investigation usually performed when analyzing the power grid concerns reliability. This is almost always the motivation that drives complex network analysis studies related to high voltage electrical infrastructures. Usually, the investigation involves evaluating the disruption behavior of the graph when its nodes or edges are removed.

Other terms to compare the various power grid studies involve more general characteristics of the network. In particular, the geographical location of the analyzed grid is responsible for topological properties due to the different morphological characteristics of different countries. Another relevant aspect deals with the layer of the power grid under investigation since differences can emerge from a topological perspective investigating the different ends in which the grid is usually partitioned: high, medium and low voltage. It is also important to have information if the type of power grid graph comes from a real network infrastructure or it is a synthetic sample extracted from blueprint models for the power grid electrical engineering literature.

The motivations to include the works are based on the thoroughness and precision of the research performed, the rigor in the application of complex network analysis methodologies and the geography of the power grid analyzed in order to cover a broad spectrum of the infrastructures realized in the different countries and identify possible differences or similarities.

A number of studies have been performed on the high voltage grid. Here we describe the most important aspects of each work under investigation. The works shown in Table 2.1 have been chosen based on the following factors: they are specifically about the power grid, they cover US, European, Chinese and Indian grids or synthetic topologies from electrical engineering literature, they have samples of different sizes and, most importantly, these are the best-known and most representative works from the complex network analysis and power grid community.

2.2.1 Power grid characteristics

Concerning the grids we consider general and non-technical aspects so to give a global idea as shown in Table 2.1. Several aspects of comparison are taken into account: the number of nodes and lines composing the grid (second and third column)[1]; the type of sample considered either a real grid or synthetic samples, for instance, coming from IEEE literature such as IEEE Bus systems (fourth column); the type of grid analyzed (fifth column) in belonging either to the transmission part (high voltage) or to the distribution part (medium and low voltage); another essential information deals with the geography of the grid (last column).

The data are in most cases extracted from real samples, that is, they represent real electric infrastructures. Other works in addition to real power grids consider synthetic grids. Most of these studies investigating synthetic samples use IEEE blueprint networks such as IEEE Bus systems, while very few concentrate only on other synthetic networks (e.g., non-IEEE models, small-world models, random graphs). Almost all samples belong to the high voltage end of the power grid. The high voltage grid contains the lines used for long range transmission to which big power plants are attached too; the only exception is our study [159] that is focused on the distribution part of the grid.

From a geographical perspective the samples are mainly localized in the United States and in Europe with some studies that consider Chinese high voltage samples, and only one that considers India; a map of the countries whose grids are analyzed is represented in Figure 2.1. Another main commonality is to treat the grid as an undirected graph where each substation or transformer represents a node and each line transporting electricity is an edge.

2.2.2 Statistical global graph properties

The main characteristics from a graph and complex network analysis perspective of the grids under analysis are summarized in Table 2.2 where we consider the *order* (N) and *size* (M) of the graph (second and third column). The average degree, computed as $\langle k \rangle = \frac{2M}{N}$, gives a general idea of how many vertexes is an average vertex connected to (fourth column). Fifth, sixth and seventh column give information about the type of statistical analysis performed on the graph, in particular,

[1]Note that the numbers in the second and third column are not the exact numbers, but they are an approximation to give the idea of the importance of the sample.

Work	Number of Nodes	Number of Lines	Sample Type	Network Type	Geography
[4]	∼14000	∼19600	Real	HV	North America
[57]	∼300	∼500	Real	HV	Italy
[42]	∼314000	N.A.	Real	HV	North America
[89]	∼4800	∼5500	Real	HV	Scandinavia
[184]	∼2700	∼3300	Real	HV	Europe
[185]	∼3000	∼3800	Real	HV	Europe
[192]	∼3000	∼3800	Real	HV	Europe
[58]	∼370	∼570	Real	HV	Italy, France and Spain
[186]	∼370	∼570	Real	HV	Italy, France and Spain
[222]	∼4900	∼6600	Real	HV	Western US
[218]	∼8500	∼13900	Synthetic and real	HV	Western US and New York State Area
[159]	∼4850	∼5300	Real	MV/LV	Netherlands
[132][2]	∼210	∼320	Synthetic and real	HV	China
[183]	N.A.	N.A.	Real	HV	Europe
[28]	300	411	Synthetic	HV	
[105]	∼6400	∼8700	Synthetic and real	HV	North America, Scandinavia and Korea
[86]	300	411	Synthetic	HV	
[219]	∼8500	∼13900	Synthetic and real	HV	Western US and New York State Area
[143]	∼30	∼13900	Synthetic and real	HV	Western US and New York State Area
[83]	∼900	∼1150	Real	HV	China
[217]	∼3200	∼7000	Synthetic and real	HV	New York State Area
[216]	∼4900	∼6600	Real	HV	Western US
[60]	∼1700	∼1800	Real	HV	China
[65]	∼39	∼46	Syntethic	HV	
[64]	∼39	∼46	Syntethic	HV	
[82]	∼2500	∼2900	Real	HV	China
[199]	∼15400	∼18400	Real	HV	North America and China
[169]	∼550	∼800	Synthetic	HV	
[179]	∼14000	∼19600	Real	HV	North America
[8]	∼90	∼120	Synthetic	HV	
[26]	∼550	∼700	Synthetic and real	HV	Italy
[87]	∼29500	∼50000	Synthetic and real	HV	North America
[176]	∼400	∼700	Synthetic		
[32]	∼900	∼1300	Synthetic and real	HV	South-East US
[38]	∼60	∼110	Real	HV	India

Table 2.1: Comparison between studies using CNA for the power grid .

the assessment of node degree distribution and betweenness distribution together with an evaluation of the path length are considered. Another term of comparison deals with the type of graph analyzed taking into account weights or not. The last two columns of the table consider the type of aim of the analysis: either an investigation of the disruption behavior of the grid, or the evaluation of the small-world properties.

Considering Table 2.2, a difference appears: the studies closer to a topological characterization use unweighted representation of the edges of the grid and consider always the node degree distribution in the analysis, since it is an important

[2]The values for nodes and lines in this table refer only to a snapshot of Shanghai power grid.

Work	Sample Order	Sample Size	Average degree	Node Degree Distribution Statistics	Betweenness Distribution Statistics	Path Length Analysis	Weighted/Unweighted Analysis	Resilience Analysis	SM all-world Investigation
[4]	~14000	~19600	~2.80	✓	✓		Unweighted	✓	
[57]	~300	~500	~3.33	✓	✓	Indirectly through efficiency metric	Weighted not based on physical properties	✓	
[42]	~314000	N.A.	N.A.	✓			Unweighted	✓	
[89]	~4800	~5500	~2.29	✓		✓	Unweighted	✓	✓
[184]	~2700	~3300	~2.44	✓			Unweighted	✓	
[185]	~3000	~3800	~2.53	✓		✓	Unweighted	✓	✓
[192]	~3000	~3800	~2.53	✓			Unweighted	✓	
[58]	~370	~570	~3.08			Indirectly through efficiency metric	Unweighted	✓	
[186]	~370	~570	~3.08	✓		✓	Unweighted	✓	
[222]	~4900	~6600	~2.69			✓	Unweighted		✓
[218]	~8500	~13900	~3.27	✓		✓	Unweighted and impedance analysis		✓
[159]	~4850	~5300	~2.18	✓	✓	✓	Both	✓	✓
[132][2]	~210	~320	~3.05			✓	Both	✓	✓
[183]	N.A.	N.A.						✓	
[28]	300	411	2.74				Both	✓	
[105]	~6400	~8700	2.72			✓	Unweighted	✓	✓
[86]	300	411	2.74	✓ (chart only)	✓		Both	✓	
[219]	~8500	~13900	~3.27	✓			Unweighted	✓	
[83]	~900	~1150	~2.55	✓	✓	✓	Weighted	✓	✓
[217]	~3200	~7000	~4.375	✓ (chart only)	✓		Weighted	✓	
[216]	~4900	~6600	~2.69				Unweighted	✓	
[60]	~1700	~1800	~2.12			✓	Both		✓
[65]	~39	~46	~2.36				Weighted	✓	
[64]	~150	~46	~2.36				Weighted	✓	
[82]	~2556	~2892	~2.26				Weighted	✓	
[199]	~15400	~18368	~2.39	✓ (for one sample only)		✓	Unweighted	✓	✓
[169]	~550	~800	~2.91			✓	Unweighted	✓	
[179]	~14000	~19600	~2.80				Weighted not based on on physical properties	✓	
[8]	~90	~120	~2.67				Weighted	✓	
[26]	~550	~700	~2.55				Weighted	✓	
[87]	~29500	~50000	~3.39				Weighted	✓	
[176]	~400	~700	~3.5			✓	Weighted	✓	✓
[32]	~900	~1300	~2.89				Unweighted	✓	
[38]	~60	~110	~3.67				Unweighted		

Table 2.2: Comparison of the main characteristics of the graphs related to power grids .

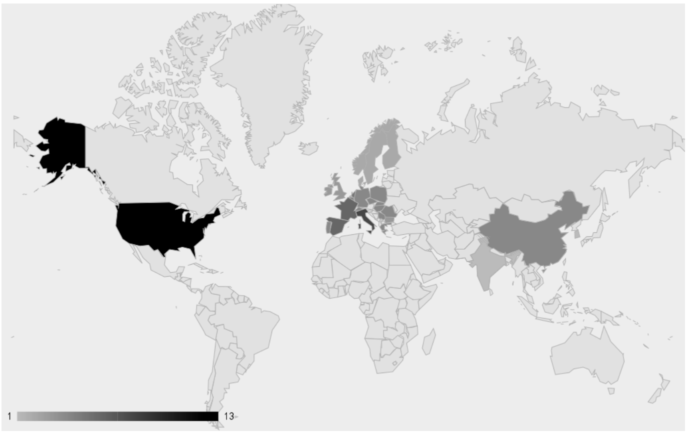


Figure 2.1: Map of the power grid infrastructure studied using CNA approach.

element to define the type of network (e.g., scale-free network). On the other hand, the studies that apply a weighted representation of the grid do not consider the node degree distribution statistics: neither considering the unweighted definition of node degree, nor using a definition that takes into account weights as proposed in [152]. This last aspect might be worth consider since the degree distribution properties of the network might change the picture of the node degree distribution in comparison with unweighted studies.

Centrality measures are not often used, exceptions are [132, 217, 28, 65] these are necessary to identify the statistical distribution of critical nodes. More attention to centrality measures, especially using weighted representation of the power grid based on the capacity or energy flows through the grid, might be beneficial in understanding the most critical nodes or lines in the power system.

Another recurring theme in the complex network analysis involving the power grid is the reliability analysis, which it actually is the main motivation that drives these kind of studies. In fact, many works were performed after major blackouts occurred, such as the North American black-out of 2003[3] or the Italian one of 2003[4] (e.g., [4, 56, 57, 42, 179]) or anyway mention blackouts as the main motivation for the work. The fragility and resilience properties of the power grid has been the major reason of concern that has determined the focus of such complex network analysis studies on the high voltage network. That is why almost all studies consider the behavior of the grid to various attacks to its nodes or edges.

[3] `http://news.bbc.co.uk/2/hi/americas/3152451.stm`
[4] `http://news.bbc.co.uk/2/hi/3146136.stm`

2.2.3 The small-world property

The small-world is a name for a property of networks where a *path* between any two nodes is usually small. It has received lots of attention starting with sociological studies [135, 204], but more recently with application of this concept and model to many more classes of networks [223, 222, 197, 108]. Among the studies analyzed, the small-world property investigation is performed by ten out of the thirty two considered. The various studies look for the satisfaction of the small-world property described by Watts and Strogatz [223, 222].

In general, there is no definitive answer to the question regarding the membership of power grid networks to the small-world group. It is indeed very specific to the samples analyzed and no general conclusion can be drawn, this seems especially true for the high voltage grids, while the medium and low voltage networks seem far from being small-world networks [159]. The path lengths are similar to those of random graphs satisfying the condition ($L \approx L_{RG}$) proposed in [223], while analyzing the clustering coefficient (γ), one sees that only for few samples the condition of clustering coefficient much higher than a random graph with the same *order* and *size* ($\gamma \gg \gamma_{RG}$) is satisfied. This justifies the statement made before stating that no general definitive answer can be given for the power grid network considering the membership to the small-world network category.

2.2.4 Node degree distribution

The degree of a node is the number of edges incident to it. However, this information is not particularly important for big graphs since keeping track of each node degree may not be manageable nor meaningful. Instead it is better to have a general idea of the statistics of the node degree for the whole graph to understand the overall tendency of the statistical distribution as a synthesis of the degree behavior of the network. In particular, its probability distribution gives us some insights of the general properties of the networks such as the likely or unlikely presence of nodes with very high degree (sometimes also referred as hubs). Table 2.3 shows the main information about the degree distribution in the works studied. The second column gives the type of cumulative node degree distribution. The third column provides the analytical function obtained in the studies by building a statistical distribution from the empirical data.

As seen in the table, the results do not completely agree on the type of the distribution followed by power grid networks, but generally they are close to an exponential decay. In the works [218, 219] presented in the table the functions $f_1(x)$

Work	Cumulative Node Degree Distribution Probability Type	Fitted Distribution
[4]	exponential	$y(x) \sim e^{-0.5x}$
[57]	exponential	$y(x) = 2.5e^{-0.55x}$
[42]	power-law	$y_1(x) = 0.84x^{-3.04}$ $y_2(x) = 0.85x^{-3.09}$
[184]	exponential	$y_1(x) \sim e^{-0.81x}$ $y_2(x) \sim e^{-0.54x}$
[185]	exponential	$y(x) \sim e^{-0.56x}$
[192]	exponential	$y(x) \sim e^{-0.61x}$
[186]	exponential or sum of exponential terms	$y_1(x) = e^{-0.18x^2}$ $y_2(x) = e^{-0.21x^2} + 0.18e^{-0.25(x-4)^2}$ $y_3(x) = 0.96e^{-0.17x^2} + 0.25e^{-0.19(x-3.9)^2}$
[218, 219]	Sum of truncated geometricaland irregular discrete terms	$y_1(x) \sim f_1(x)$ $y_2(x) \sim f_2(x)$
[159]	power-law (unweighted) and sum of exponential terms (weighted)	$y_1(x) \sim x^{-1.49}$ $y_2(x) \sim 0.15e^{-21.47x} + 0.84e^{-0.49x}$
[83]	exponential	$y_1(x) \sim e^{-0.65x}$ $y_2(x) \sim e^{-0.58x}$
[199]	exponential	$y(x) \sim e^{-0.5x}$

Table 2.3: Comparison of the node degree cumulative distribution probability functions.

and $f_2(x)$ are not reported in the table for size reason, but in footnote.[3] For studies concerning multiple samples (i.e., [184, 185, 159]) averages between all samples, or particular significant samples have been chosen among the many available.

In general the various studies focusing on the high voltage grids agree on a statistical distribution for node degree that follows an exponential (or exponential based) distribution with characteristic parameters of the exponential curve that depend on the specific grid. While high voltage grids have been quite extensively analyzed, the medium and low voltage grids have not found much attention so far and a deeper and wider investigation needs to be performed in different geographies since the only study (i.e., [159]) is representative of the Northern part of the Netherlands. In addition, the distribution grid will be the section of the power grid mostly impacted by innovations in future power systems i.e., smart grid technology [31].

Very few studies in the literature have shown results of an accurate methods such as the one based on Kolmogorov-Smirnov method, shown by Clauset *et al.* [47]. As remarked in [47], many distributions that with a naive fitting analysis might be considered power-laws at first sight do not prove to belong to such a category when a more rigorous fit test is adopted.

[3] $f_1(x) = \sum_{x_i<x} 0.2269(0.7731)^{x_i} * \{0.4875\delta(1), 0.2700\delta(2), 0.2425\delta(3)\} \quad x_i = 1, 2, \ldots 34$
$f_2(x) = \sum_{x_i<x} 0.4085(0.5916)^{x_i} * \{0.3545\delta(1), 0.4499\delta(2), 0.1956\delta(3)\} \quad x_i = 1, 2, \ldots 16$
The $*$ symbol is here to be considered as the convolution operator and the δ is the Dirac delta function.

2.2.5 Betweenness distribution

Betweenness is an important measure to assess how a node is central in a network. This metric in fact computes how many shortest paths traverse a node, therefore giving an information of the importance of the node in the path efficiency. The main characteristics of betweenness studies are summarized in Table 2.4 where the second column shows the type of followed distribution, while the analytical function is represented in the third column. This metric is computed by only five studies (i.e., [4, 57, 159, 83, 217]). Although the studies that perform this type of analysis are only few, one can see that there is a tendency for the high voltage network to have a betweenness distribution close to a power-law. For the medium and low voltage the situation is less clear: some samples analyzed in [159] follow an exponential decay, especially the low voltage ones, while other, usually the bigger belonging to the medium voltage, follow a power-law.

Work	Cumulative Betweenness Distribution Probability Type	Fitted Distribution
[4]	power-law	$y(x) \sim (2500 + x)^{-0.7}$
[57]	power-law	$y(x) \sim 10000(785 + x)^{-1.44}$
[159]	power-law and exponential	$y_1(x) \sim x^{-1.18}$ $y_2(x) \sim 0.68e^{-6.8 \cdot 10^{-4}x}$
[83]	power-law	$y(x) \sim x_1^{-1.71}$ $y(x) \sim x_2^{-1.48}$

Table 2.4: Comparison of the betweenness cumulative distribution probability functions.

Power-law seems the dominating rule for betweenness probability distribution even if few studies consider this statistical property of graphs. To draw a general definitive conclusion regarding this property for the power grid more studies are required. Another aspect to be considered which has not received much attention so far is the study of betweenness statistics in weighted grid models or when power flows are considered instead of the pure topological analysis.

2.2.6 Remarks

Complex network analysis studies and their comparison show how important properties of a real system such as the power grid can be studied using graph modeling tools and which conclusions about the reliability of the infrastructure can be drawn. CNA proves to be a good set of tools that provide, although without dealing with the details and complexities of the electrical properties in the case of the power grid, a comprehensive and general understanding of the properties that characterize a network. We see an interesting trend in the various works analyzed, that is, the research towards more complex representation of the properties of the network than a simple graph. In fact, although complex network analysis can help in understanding the foundational properties of the network aspects of the power gird, it is always worth to remember that the power grid is subject to the law of physics and the principles of electrotechnics. From the initial studies (e.g., [222, 4, 42, 89, 199]) considering the power grid just as an undirected graph without any property (i.e., weight) on edges and with no characterization of the nodes, more recent studies take into account the electrical properties of the power grid system. Of course the aim of these later studies is always to provide, anyway, a simplification of the highly complex power system. However, they add those essential parameter to better model the grid characteristics: impedance parameters associated to the transmission lines, power limits supported by the substations (i.e., nodes) such as for instance in [132, 26, 87, 82, 217]. A more detailed description of the power grid under investigation (i.e., weighted graph representation) enables to better understand the dynamics guiding the power grid with a mixed approach: both preserving the idea of the complex network analysis of having a general and statistical behavior of the overall power grid, and, on the other hand, to take into account the physical/electrical properties essential to characterize the power grid. Latest results in [87, 26, 8] show a better agreement to real power grid behavior of the models that take into account physical parameters, compared to the pure topological analysis, and the observed behaviors and critical points in real power systems thus justifying this *enhanced* complex network analysis approach.

We also see another gap in the current scientific work concerning the power grid and complex network analysis techniques: the applicability to the real power systems of the results obtained by the network analysis. Especially, the findings of the studies that consider the vulnerability and cascading effects of the power grid need to be confirmed by the transmission and distribution operators. We remark here that basically all the works lack the cross-check of the theoretical results with the experience on the field.

A noteworthy general aspect is the role that complex network analysis has for the power grid infrastructure vulnerability analysis: CNA does not want to substitute the traditional approaches to power systems resilience and safety analysis since they have proved extremely successful in governing and managing the electrical system with only occasional highly disruptive events. CNA techniques applied to the power grid world represent a simplification of all the complexities governing the power systems, but such approach can anyway be helpful to give a general vision that can help in identifying quickly and in a simple way critical spots or aspects of the power grid which then may be further and deeply analyzed with traditional electrical engineering tools.

2.3 Power Grid, Network Design and Evolution

Network dynamics consider the change over time of a network. It consists basically on how new nodes and links are added or the current ones are removed. We look at evolution with specific focus on link formation concerning complex networks. We then draw the attention on complex network analysis studies to evolve the high voltage power grid and then we focus on traditional electrical engineering power lines design and adaptation.

2.3.1 Evolution of complex networks

The study of the dynamics evolution of networks by using complex network analysis techniques has been widely investigated. The focus is mainly to assess and analyze existing networks that have evolved over time and understand which principle is underlying such evolution. More in general, the evolution is considered in defining abstract algorithms that can be used to build networks that mimic the same behavior of natural or technical networks. For a general introduction to complex network analysis generation techniques, we refer to extensive works such as [39, 151]. The work of Liben-Nowell and Kleinberg [121] discusses link prediction. The authors consider the evolution of the ArXiv library co-authorship network (i.e., a social network) and define several quantitative metrics to infer the future links to be formed in the network (i.e., new co-authored publications). The metrics that the authors use for predictions are based on evaluating the paths, the neighborhood of nodes, rank of the matrix representing the graph, and the clustering. Compared to a random prediction of new links the methods used in the paper are far better; in fact the random predictions is correct less than 1% of the times, while on some topics of the ArXiv library the predictions reach a correct outcome 50% of the time. The ideas used in the paper and the metrics used in the prediction of new connections between nodes prove valuable for the co-authorship social network. Even if at the abstract level, networks are the same (i.e., graphs) it is difficult to apply these metrics and prediction strategies on a real infrastructure such as the power grid. In fact, many of the metrics considered in the paper are inspired by the underlying social aspects of the co-author relationship such as neighbors in common. This aspect is reinforced by the different results that the same metrics used on different topics of the ArXiv library sometimes provide. Given the small average node degree of the power grid networks, the absence of samples at different time stamps, it is difficult to apply the metrics and methods of the paper to distribution networks and expect meaningful and interesting results.

The survey of Lü *et al.* [125] provides a good overview on link prediction. The authors propose several methods and algorithms based on similarity properties of the nodes at local level (e.g., the neighborhood of a node), global similarity (i.e., over the whole network), and quasi-local that combines the best of the two approaches. They also propose other strategies such as probabilistic models to evaluate the likelihood of the formation of a link. One of the networks that is considered for link prediction in [125] is the Western U.S. power grid (data already used in [222]). In this case the methods that give more stress on the local interactions between nodes perform better in the prediction than other approaches. The authors also state that for infrastructural networks the geography plays an important role where the long distant connections are discouraged and very rare. The paper describes three main applications of link prediction in the field of CNA: reconstruction of networks, evolution mechanism of networks, and classification of partially labeled networks. We consider these prediction mechanisms valuable for networks with rich neighborhood structures where differences and intersection between neighborhoods are meaningful (e.g., social networks) or where the discovery of links is particularly costly or long in time and efforts (e.g., protein interactions). The current power grid networks have low average node degree and the structure is usually close to radial, especially at medium and low voltage level, thus we consider the predictions less interesting. However, these techniques and link predictions might be more applicable in the further evolution and prediction of new links when the networks are more rich in connectivity. In addition, some techniques for link prediction would benefit from additional information available to characterize the nodes, i.e., the number of prosumers that are connected to a node (i.e., substation) and the available production power.

The concept of network evolution of a technological infrastructure is considered in [177] where a model for the evolution of the Internet Autonomous System (AS) is provided. As most works (e.g., [61, 12]), this one focuses on creating a model of the existing evolution of the Internet rather than proposing new ways of evolving networks. The interesting aspect on the modeling of the Internet is the better results of the proposed model defined as Parallel Addition and Rewiring Growth (PARG) compared to other models in the literature to capture specific features of the AS topology. In particular, the model is able to capture the dissortativeness of hubs (i.e., the most connected nodes in the network) and linear local assortativeness of nodes. The model proposed proves successful, however one must recognize that the AS connections are mainly logical and are due more to business and contractual agreements than physical interconnections, therefore missing part of the problem that is essential in the case of the power grid topology.

2.3.2 Complex network analysis for power grid design

Few studies in the complex network analysis landscape consider the possibility of using the insight gained through the analysis to help the design. These few cases consider the addition of lines in the network to assess the increase in the reliability of the entire power grid. Examples are the study of the Italian high voltage grid [58] and the study of improvement by line addition in Italian, French and Spanish grids [186]. Also Holmgren [89] uses complex network analysis to understand which grid improvement strategies are most beneficial. He shows the different improvements of typical complex network analysis metrics (e.g., path length, average degree, clustering coefficient, network connectivity) although in a very simple small graph (less than 10 nodes) when different edges and nodes are added to the network. Broader is the work of Mei *et al.* [132] where a self-evolution process of the high voltage grid is studied with complex network analysis methodologies. The model for power grid expansion considers an evolution of the network where power plants and substations are connected in a "local-world" topology through new transmission lines; overall the power grid reaches in its evolution the small-world topology after few-steps of the expansion process. Wang *et al.* [218, 220] study the power grid to understand the kind of communication system needed to support the decentralized control. The analyses aim at generating samples using random topologies based on uniform and Poisson probability distributions and a random topology with small-world network features. The simulation results are compared to the real samples of U.S. power grid and synthetic reference models belonging to the IEEE literature. These works also investigate the property of the physical impedance to assign to the generated grid samples.

Complex network analysis is not generally used as a design tool to propose new topologies for the future smart grid as we use in this work where we also assess the benefits in terms of economical improvement.

2.3.3 Power engineering approach for power lines design

Traditionally power system engineers adopt techniques from electrotechnics, which model deterministically the power lines and their characteristics, these might include the exploitation of graph theory principles [214, 54]. The traditional techniques applied by power engineers involve the individuation of an objective function representing the cost of the power flow along a certain line which is then subject to physical and energy balance. The problem translates into an optimization problem. Since long time the models are applied both for the high voltage [74, 116] and the medium and low voltage planning [214, 54]. Not only optimization, but also expert systems [126] are developed to help in the process of designing grounding stations based on physical requirements as well as heuristic approaches based on engineering experience. The substation grounding issue is approached as an optimization problem of construction and conductor costs subject to the constraints of technical and safety parameters, its solution is investigated through a random walk search algorithm [79]. In [71] a pragmatic approach using sensitivity analysis is applied to a linear model of load flow related to various overloading situations and a contingency analysis (N-1 and N-2 redundancy conditions) is performed with different grades of uncertainty in medium and long term scenarios.

In the practice, the planning and expansion problem is even more complex since it implies power plants, transmission lines, substations and the distribution grid. In [81] all these aspects are assessed separately and several challenges appear. For instance in the planning of a high voltage overhead transmission line, specific clearance code must be followed and not only load is a key element, but also topography and weather/climatic (above all wind and ice) conditions play an import role in the planning of the infrastructure. For substation planning the authors of [81] emphasize, in addition to the need for upgrading the grid (e.g., load growth, system stability) and budgeting factors, the multidisciplinary aspects that come from environmental, civil, electrical and communications engineering. A more general approach is proposed in [81] to deal with power system planning; it might be regarded as a multi-objective (e.g., economics, environment, feasibility, safety) decision problem requiring the tools typical of decision analysis [100].

The works mentioned so far take into account mainly the high voltage end of the grid while not least important is the distribution grid, especially in the vision of the future electrical system where the end user plays a vital role. The integrated planning of medium and low voltage networks is tackled by Paiva *et al.* [170] who emphasize the need of considering the two networks together to obtain a sensible optimal planning. The problem is modeled as a mixed integer-linear programming one considering an objective function for investment, maintenance, operation, and losses costs that need to be minimized satisfying the constraints of energy balance and equipment physical limits.

Even more challenges to electrical system planning are posed by the change in the energy landscape with several companies running different aspects of the business (generation, transmission, distribution) and by the increasing of prosumers and renewable energy resources since these players have less control on each other compared to a fully integrated electricity system and they also have different objectives. In addition, accommodating more players in the wholesale market transmission expansion should follow (as it is already for generation) a market based approach i.e., the demand forces of the market and its forecast should trigger the expansion of the grid [30]. The same considerations regarding the need for a different approach in planning in a deregulated market are expressed in [190] where the optimization of an objective function in the market environment is applied. Another method to evaluate transmission expansion plan takes into account the probability reliability criteria of Loss Of Load Expectation (LOLE). In particular, in [45] an objective function is proposed that takes into account the cost of constructing a transmission line between all buses involved. The designed line is then subject to constrains in peak load demand satisfaction and a certain level of LOLE that the line should not outrun.

In the smart grid framework the planning techniques might be revised especially for the distribution grid which is the segment that is likely to face the greatest changes due to the presence of Advanced Metering Infrastructure (AMI) (i.e., bidirectional intelligent digital meters at customer location) and Distribution Automation (DA) (i.e., feeders can be monitored, controlled in automated way through two-way communication). In addition, the medium and low voltage grid is no longer a layer where only energy is consumed, but distributed energy generation facilities (small-scale photovoltaic systems and small-wind turbines) will be attached to this segment of the grid; altogether these elements are likely to reshape the way planning for medium and low voltage are realized [31].

2.3.4 Power lines adaptation

A power engineering problem relevant to the scope of this work is that of distribution grid topological reconfiguration. The early works on the topic of Merlin and Back considered the problem with heuristic techniques and algorithms (e.g., branch and bound) to reconfigure the switches of the distribution infrastructure [134]. Later works of Cinvilar *et al.* [46] provide computationally tractable methods to modify the topology of the distribution grid by opening/closing switches to reduce losses. The authors provide a simplified formula to estimate the reduction in losses between the situation before and after the reconfiguration in the topology. Baran and Wu [17] have also proposed the idea of minimizing the losses and guaranteeing the load balancing. Basically, the problem deals with finding the minimal spanning tree of the network, since the radial configuration needs to be preserved, that minimizes the objective function (i.e., system losses) while satisfying the constraints on voltage, capacity of lines and transformers, and reliability. For this problem, Baran *et al.* use a simplification of the power flow equations to compute the power flow in the network to be optimized.

Recently, new interest has grown concerning the reconfiguration problem together with islanded grids and micro-generation plants. The work of Ramesh *et al.* focuses on the minimization of losses in the distribution grid [180]. The authors provide several options that have been simulated and tested on the field to reduce losses on real and reactive power. The three solutions proposed concern: distributed generation, capacitor placement and restructuring of feeders. The first proposal consist in placing local energy generation closer to the end users. In certain buses of the simulated IEEE 37 Bus, they show loss reduction of up to 9 MW. The second solution proposes the installation at optimal locations of capacitors. The third proposed technique concerns the restructuring of the topology of the network. The solution proposed is basically a reconfiguration of the opening and closing of the switches of the distribution network to maintain the network radial and satisfy its load requirements. This approach although valuable from the practical point of view, lacks the vision in imagining the future energy scenario and there is no investigation of which topology is best, or according which principles the new pieces of infrastructure (e.g., cables) should be interconnected. The aspect of network reconfiguration of power networks to achieve a minimization of losses is the topic of [144]. The authors state that the only way to improve efficiency is by altering the topology. They focus on two aspects: (i) finding an optimal switching scheme of the switches connecting the lines and minimize losses, (ii) adding lines. The case study used is a IEEE 14 Bus where three more lines are added. The authors do not give any details why only three lines are considered and the motivation to choose specific nodes to attach the new lines. In addition, the only argument provided concerns the economic aspect of adding only the three lines, but no quantitative evidence is provided. In [191] the problem of network reconfiguration in a smart grid environment is addressed. The idea of the authors is to reconfigure the topology of the network by operating switches in order to minimize the overload in the branches of the network. The scenario they consider takes also into account the higher penetration rate that distributed generation will have in the future. The numerical evaluation of the proposed genetic algorithm to reconfigure the network is realized on a simple 33 Bus network where the system configuration problem is solved to minimize an objective function (i.e., losses) and without violating the voltage and current constraints on the lines. Even in that work, the importance of topology is claimed and the benefits in terms of reduced losses are shown. However, no additional lines or investment of the distribution grid are proposed and the benefit come just from the reconfiguration. No motivation or order of magnitude of the avoided investment are provided and the effects of topological changes available in the switching are anyway limited. The work of Xiaodan *et al.* [226] is in the spirit of studying the reconfiguration in a micro-grid environment with micro-generation plant based on renewables. Basically, the problem is considered as two optimization sub-problems: one related to determine the capacity of each island of the grid that has micro-generation capabilities and the other is a problem of reconfiguration of the distribution grid with the objective of minimizing the power losses. The optimization algorithm is validated against two test grids: the IEEE 33 Bus and the PG&E 69 node system. The reconfiguration for the system in an island and micro-generation uses techniques and approaches similar to traditional reconfiguration of distribution systems. The problem considered is an optimization problem and the authors constrain the distribution grid to be always operated in a radial configuration. Usually, the reconfiguration problem boils down to the definition of an objective function to minimize the losses of the system and to establish constraints to satisfy the load. The function is then solved resorting to some heuristic.

2.4 Smart Grid Services and Applications

The smart grid promises to bring changes to the electrical systems by promoting a massive use of ICT technologies to enable the various actors in the system to interact and create new services. Therefore, the ICT aspects concerning the smart grid have received even more interest than changes in the infrastructure so far. We summarize the main features of the literature concerning smart grid software applications and architectures, service orientation and energy, and energy market-oriented solutions.

2.4.1 Smart grid software applications and architectures

The smart grid has grown in popularity in the past few years and several approaches and architectures have been proposed for the information services related to the smart grid for the smart home. Most of the works are architecture proposal with interfaces to services that are implemented abstractly or based on historical energy data, thus, missing to model the real-time dynamics of the system.

In [188], the authors simulate the effects of aggregate households appliances and electric vehicles to act as balancing devices for solar installations. They show the feasibility of reducing the amount of spinning reserves and the benefits in terms of reduced CO_2 emissions, considering several types of solar plants and scenarios of electric vehicles penetration. The work is entirely a simulation effort with no real services involved. Another simulation tool for demand-response is proposed in [215]. The goal is to analyze at a very detailed level (bus voltages, power flows, etc.) a control strategy to accommodate the wind power of a turbine and its fluctuations (0.5 MW of power) in a small/medium-scale community (650 buildings). The model shows that by applying a control strategy to the heath pumps in the buildings one can provide a considerable smoothing in the variations of power in the slack bus of the simulated IEEE 13 bus system. The presented system is a complex MATLAB simulation environment and deals with the electrical problem and not with the ICT or Internet-based services to enable the demand-response in practice. The importance of services in the future smart grid is emphasized by Karnouskos who considers the services as the only way to let the various actors in the smart grid panorama interact with each other [95]. It is more a vision/position paper than an analysis, and the work does not propose specific services, technologies or implementations to be used. The work of Strobbe *et al.* [196] provides an architecture for the future smart grid that is built on JAVA-based standard protocols. The services provided are mainly proposed to engage and raise the awareness of the user to energy conservation and peak reduction by having information of prices. The architecture presents interesting aspects such as the ease of extension and the energy price individuation via the Belgian power exchange trading market. However, there is no mention of the integration of renewable sources and energy forecast in the architecture, that are essential building blocks of the future smart grid. In addition, the work does not consider any automation process towards appliances, but just the user interaction with an home display that provides consumption and tariff information. Another example of an architecture based on services is given in [96]. Although the amount of services proposed for the future smart grid is considerable (energy monitoring, prediction, management, optimization, billing, and brokerage) and well described also in the technical details of the architecture (e.g., representational state transfer based), there is basically no description of the implementation of such services. There are no details of data sources in order to implement such architecture in practice. Another architecture for the future grid is proposed in [211]. The authors consider a service architecture that can easily plug-in new services. The architecture can host smart applications that are able to interact and control devices, and act as a contact point with the services provided by service-providers related to energy (e.g., real-time pricing, remote device control). The simulation of the smart grid is obtained through the OMNeT++ framework where the information about topology, user consumption and scenario parameters are provided. The architecture misses the interaction with real information and data that are the key to provide more concrete and valuable results to the mere simulation exercise. The model presented in [33] focuses on simulating a multi-agent system where each agent is an energy consumer or producer. The simulation is based on historical demand patterns, while in the energy generation patterns, context information for solar and wind production are taken into account through the interaction with Web services. Although the approach is interesting and uses real data, it lacks an important ingredient of the smart grid that is dynamic pricing. Moreover, there is no mention of energy forecast that is essential in the schedule operations of home appliances, and the time granularity considered in the work is quite long (three hours) compared to the dynamics of the smart grid. A residential energy management system is described in [182]. In that paper, the demand-response is well described, there are several devices and scenarios taken into account (dishwasher, dryer, electric vehicle, refrigerator, etc.), and a return on investment analysis is performed. Though, there are some limitations. First, it is only a simulation and there is no interaction with real equipment or real services providing consumption information; second, there is no tariff differentiation and the demand-response is only available in a time window during the day; third, distributed energy generation through renewables is not mentioned. Web services are the key components to interact in a smart grid-enabled home as suggested in [103] and [9]. These works, however, are mostly an exercise in Web service interaction and simulation of communication (representing appliances and sensors) rather than a real set of services or components that can be applied in a smart grid test solution. The data related to energy consumed in the home environment is the U.S. average, the energy extractable by solar panels and wind turbine are static and with no variation according to environmental condition, moreover only peak and off-peak tariffs are considered. Such approach of simulating the smart grid is quite simple and lacks in the dynamism of the smart grid components (i.e., energy prices, renewable energy production) which are all the crucial aspects of the smart grid. Considering the broader aspect of the whole interactions between the many actors of the smart grid, service-oriented architectures (SOAs) have been proposed in several works to be the good candidates to solve interoperability issues [161, 198, 158].

In summary, we note that effort has been spent in the simulation of smart grid aspects; while there is a gap in the investigation of how the future smart grid services should be and how they should work. Even more problematic is that the data and information used to simulate the smart grid (e.g., dynamic pricing, usage of renewable energy) are usually based on limited historical data, often averaged out, rather than based on real-time actual data.

2.4.2 Service-oriented architectures for energy

Here we provide a survey of the most relevant research efforts with respect to the vision of an open grid and its relation with service-oriented architectures.

SOAs are a way of building scalable and interoperable distributed systems. The SOA paradigm is based on the concept of loosely coupled entities, called services. Service providers perform operations that can be accessed by service consumers. Provider entities publish the services together with the requirements for the access; the service consumers perform discovery to have information on the available services and the way to contact them. Once this information is obtained the service consumer can directly bind to the service provider and exploit it. A successful implementation of the SOA paradigm is through the use of the eXtensible Markup Language (XML) technology known as Web services. Web services have proven essential in providing interoperability in heterogeneous systems environments. The energy sector with many different systems is an ideal candidate to implement the SOA paradigm especially considering the future view of even more heterogeneous interoperability required by the smart grid.

The need for integration of the various actors and the corresponding information systems is something that is not new and that has been under consideration since the first signals and attempts to unbundle the energy sector. Back in 1995 Dahlfords and Pilling [59], exposed the necessity of having a dynamic and flexible information system. They foresaw the need of companies to be able to interact with many more players than they were traditionally working with. The suggestion that the authors propose in the paper is to have a two-way communication interaction with the metering apparatus along the grid. This concept is still valid since it is one of the key aspects of the modern concept of the smart grid. Notable is the vision stated in the paper: "The energy market changes from a Producer *push-market* to a Customer *pull-market*", which today we could update with what we could call a *Prosumer pull-push-market*.

The same integration theme is highlighted by Becker *et al.* [18]. They stress the need for flexible and shared information systems to let the utilities operate more efficiently in the new deregulated energy landscape. The integration is not something new according to the authors, but previous attempts in the energy sector were not at all handy. In fact, sometimes the integration were made manually, or made in a time consuming way such as point-to-point techniques that require great efforts. The solution they mention, that has also become part of standards such as those proposed by the International Electrotechnical Commission (IEC) (i.e., IEC-61970), is to use the Common Information Model (CIM) representation for devices and objects in the energy domain and use a Generic Interface Definition (GID) to expose the APIs that can be accessed. The solution proposed for the integration of loosely-coupled applications (e.g., enterprise resource planning, supervisory control and data acquisition, customer relationship management, energy management system, Energy Trading) is based on a message oriented middleware and a message broker that together enable the creation of a message bus in which the applications send and receive data; this can be seen as a predecessor of modern SOAs. An added value is given by self-describing messages for instance those encoded using XML language.

The data integration issue between the various actors and the benefit obtained by the adoption of SOA are addressed in [101]. An implementation of a Web service SOA for the energy management system/supervisory control and data acquisition systems based on top of IEC standards gives several advantages combining the power system oriented aspects (IEC defined) and the flexibility and the spread of Web services that provide easier integration with other companies, reuse of existing infrastructure and smooth development of the Web service environment [133, 127, 78]. Web service technology is also used as the communication layer to enable the interaction between all the real-time agents that at different levels are present in the components of a modular and scalable architecture [140]. This design is suitable for the vision of a smart grid composed of a huge quantity of devices communicating electricity-related data through the Internet.

SOA is seen as the glue for the new smart grid that can enable both intra-enterprise interactions and can be even and more present in the inter-enterprise interactions that characterize even more the smart grid domain. This is the idea and the approach presented in [178] where the inter-enterprise information exchange interactions modeled in the enterprise architecture framework are linked with the inter-enterprise data exchange (which are based on the IEC standards) by the definition of an ontology that can map inter and intra-enterprise domains.

Although all these works are interesting and point out specific aspects of the interaction between energy systems, they tend to miss the vision and the evolution that might arise with a smart grid infrastructure. In fact SOA in these works tend to focus on data integration inside a company or enable interoperability between different companies in the energy business value chain, but no one points out SOA characteristics for the incoming next generation grid .

Service-oriented architecture and smart grid

As we have mentioned before, the smart grid is also referred to as the *Energy Internet* [73, 206]. In such a parallelism, to enable the interoperability between a multitude of actors and applications in various works, SOAs are indicated as central to the smart grid. Especially at the household level, an SOA approach enables easy interaction between heterogeneous devices. Warmer *et al.* [221] stress how a service based architecture can be beneficial in a smart home in the new paradigm of smart grid. They see the Internet and Web services as the key to enable the interaction between the house with its smart devices and the supply companies and electricity Distribution Systems Operators (DSOs) to exchange supply bids and demand-response related functionalities. The authors call for an ontology for the smart grid domain so that the different actors can seamlessly interact with a common language. The issue related to ontology is addressed by Considine [50] who remarks the necessity of an ontology for the smart grid, actually referred to as *service-oriented grid*.

Collaboration between future smart grid objects, appliances and devices in order to achieve better energy management and efficiency is the idea of [94]. The author envisions a collaboration between different entities such as energy resources, energy marketplaces, enterprises and energy providers through Web services, since they enable flexible integration without the problems due to implementation details. In Karnouskos's vision, a device of the smart grid will be *SOA-ready* exposing in a standard way the services it can provide, and at the same time it will be able to dynamically discover services of other devices through Web Service Dynamic Discovery specifications. One of this devices is the smart meter which acts as service provider for an enhanced business process in which the meter can not only provide real-time information, but also take decisions related to energy usage and consumption interacting with other services on the Internet [98].

Cox and Considine stress how collaboration is the essence of the smart grid and only through an interaction between the many actors involved an effective implementation of the smart grid may be realized [53]. Among the requirements the authors identify as fundamental (that also appear in the NIST and Grid Wise Architecture Council stack) some aspects such as transparency, composition, extensibility and loose coupling are presented, which are also basics for SOAs. The authors also identify the standards for information exchange to be used in the smart grid for some aspects such as scheduling and time functions, weather information, device discovery and market interactions. All these elements fit in a SOA framework.

2.4.3 Agent interactions in the energy market and the smart grid

Even before the advent of the smart grid ideas, researchers have started to investigate agent-based platform to deal with the new energy market challenges that the unbundling process has brought to energy utilities companies and to energy system operators. Especially, the financial aspects of an unbundled energy market and their relation to technology have attracted research attention, e.g., [195]. The Web has been considered the appropriate solution to accommodate the requirements of an unbundled market to be efficient, transparent, and equitable [128]. Agents-based platforms have been proposed initially in the literature to model and test the approach of a multi-player market and the technology to support it. This is the flavor of the work of Lai *et al.* [113], where the issues of a modern unbundled market with several actors can be properly tackled with agents given their characteristics of autonomy, social ability, reactivity and proactiveness. In [113] the authors already saw the agents as the entities that would interact on a fully electronic market place. A more conservative approach was proposed in [70] where an agent-based architecture had the purpose of testing the deregulated market in Croatia. The importance of the market and trading of energy has become so important that power analysis tools are now enriched with energy market simulation functionality to provide a richer picture so that integrated simulations of both the aspects can be performed [122]. Another platform based on agents is presented in [92]. The authors consider an evolved energy market where the contracts are negotiated autonomously, but with the possibility of human intervention. The agents have strategies in the contract negotiation and can also learn to adapt during the various rounds. Other studies have focused on the proposal, test and evaluation of business strategies to apply to this new kind of market to satisfy equilibrium [201]. Sometimes, interoperability requirements in unbundled markets are addresses, but then the implementation follows agent specific communication languages (e.g., Knowledge Query and Manipulation Language, Protocol Operational Semantics) [113, 92]. To the best of our knowledge a study of where the agents reside, how they interact, how these architectures would scale are not addressed in significant detail. In [193] an energy market operation system is proposed. The architecture described, although based on Web services, does not completely clarify what are the services available for the market participants to interact with. Other solutions have been realized [10] to simulate different types of commodity markets (e.g., cotton, corn, electricity), where general services and interoperability requirements for a SOA representation are described. In [209] the focus is modeling both the electricity and gas energy systems in terms of infrastructure and market characterization. The authors focus on realizing an integrated model of these energy systems and markets with agents representing the main actors involved. The work is more a modeling effort using agents than a solution where agents can interact and trade energy and evaluate the market dynamics.

In the smart grid context the ideas of having agents to deal with the interactions on the market is one of the cornerstone of the approach. The idea of an agent-based market where agents trade energy is the core of [111]. Agents are characterized by a demand (and supply) that changes with the varying price. The goal is to have a bidding system between all parties involved in order to achieve an equilibrium (demand and supply) at a certain price for energy. To realize this agent mechanism called PowerMatcher, several agents are involved such as local agents that bid energy (demand or supply), auctioneer agents that deal with the price formation process, agents that aggregate the local demand and supply and bring it to an higher level of interaction (e.g., cluster). Agents are considered as the key element in the new paradigm of distributed generation and prosumers [212]. The vision is of a fully automated trading mechanism where the price varies as a function of the availability of energy and the congestion on the network. The paper stresses that the proposed market is different than what happens in the wholesale market since the agents do not reveal their reserve price and their pattern in generation/consumption. Another aspect that is emphasized is the use of micro-storage and the possibility of controlling the storage and energy sell/purchase on the market. The authors warn that agents that react with their storage to a real-time price of energy could create problems in the system, thus proposing for stability and safety purposes that the price of energy storage is known and fixed for the whole day. Agents have also been used to realize simulation of smart grid in a smart city environment [97]. In the paper, agents do not only represent the power stations, but more in general households, appliances and electric vehicles are simulated. The goal in the work is not to simulate the auction process and the negotiations, rather it is to consider the interactions between entities, and test the real-time response of agents and the possibility of interacting with devices as if they were managed in a demand-response fashion.

Agents are a very good modeling technique to deal with energy markets in general and they are even more adequate in the distributed panorama of the smart grid. The very inner properties of agents such as autonomy, flexibility, reactivity and proactivity in their behaviors make them perfect to deal with an auction system where interactions are required and agents have to adapt to the new conditions of the rapidly changing market pursuing the goal of the user. Also the social ability of the agents, if properly programmed, can help in achieving solutions that help in maintaining the balancing and safety in the network and socially acceptable energy prices.

Chapter 3

Network Models for the Smart Grid

As introduced in Chapter 1, the power grid is one of the core elements of any power system and it is the infrastructure that enables the transport and distribution of electrical energy from the producing units to the end users. Following the vision and the changes that the smart grid is likely to enable, we consider and focus our attention to the distribution grid. The distribution grid is the terminal part of the power grid infrastructure that reaches the end user; we consider that this part of the power network will experience to a great extent the changes supposed by the implementation of the smart grid.

The scientific literature exploiting complex network analysis (CNA) has only focused on the high voltage end of the power grid, though with the new vision of the smart grid this is no longer enough. Our aim is to study the topological characteristics of the distribution grid, find the affinity to well-known literature models and propose designs for the new grid. In addition, we consider economic aspects for the distribution grid, that is the effect of topology on the price of electricity to understand if current different distribution networks have different parameters influencing the electricity distribution costs. This investigation of both the topology and the relationship of the latter with economic aspects is studied considering a set of samples belonging to the Dutch distribution grid.

3.1 The Dutch Medium and Low Voltage Grid

We focus on the medium and low voltage power grid networks of (Northern) Netherlands, the former has voltage $10kV \leqslant V_{MV} \leqslant 20kV$, whereas the latter has voltage $V_{LV} \leqslant 10kV$. This choice is dictated by the fact that the Netherlands has a modern infrastructure with one the highest availability in Europe with the electrical system being 99.99486% of the time available with just an average of 27 minutes downtime per year per customer [149]. The Dutch high voltage power grid is owned and managed by one player, Tennet, while the lower layers are partitioned geographically among fourteen companies. The partition of the territory among energy distribution companies is shown in Figure 3.1. In the figure each zone has an identifier and a distribution provider: 1) RENDO Netwerken, 2) Cogas Infra en Beheer, 3) Liander (former Continuon Netbeheer), 6) Stedin (former Eneco), 7) Westland Infra, 8) ONS Netbeheer (now Stedin), 9) DELTA Netwerkbedrijf, 12) NRE Netwerk, 13) Enexis (former Essent Netwerk), 14) InfraMosane (now Enexis).

Figure 3.1: Distribution companies over the Netherlands. *(Source: www.energieleveranciers.nl)*

The grid information used in this study is provided courtesy of Enexis B.V., the distribution operator of Northern Netherlands and other regions of the country. The provided data includes information about the transformers in the grid together with the distribution substations. The data set also provides information about the distribution lines used to connect substations containing the length of cable and other interesting physical properties (e.g., resistance, capacity, voltage).

For the sake of precision, we define the notion of a (Weighted) Power Grid graph given the information available from the network samples.

- All the substations and transformers are considered equal and are represented as nodes of the graph.
- The cables connecting the substations are considered equal despite the differences in voltages and current carried and their physical properties, and thus modeled as unweighted edges in the graph.
- For the data samples that present disconnected components, each component is treated as a distinct graph.

	PRESENT STUDY						RANDOM GRAPH		
ID	Order	Size	Avg. d	APL	CPL	γ	APL	CPL	γ
1	17	18	2.118	3.398	3.313	0.00000	1.427	1.688	0.13726
2	15	16	2.133	3.086	3.000	0.00000	2.319	2.358	0.00000
3	24	23	2.087	4.499	4.228	0.00000	3.127	3.091	0.05508
4	30	29	1.933	4.545	4.449	0.00000	1.860	2.242	0.05778
5	188	191	2.032	17.726	17.878	0.00000	3.846	4.345	0.00532
6	10	9	1.800	2.423	2.223	0.00000	0.978	1.167	0.26667
7	63	62	1.968	5.204	5.404	0.00000	2.514	2.904	0.03175
8	28	27	1.929	4.784	5.000	0.00000	2.553	2.945	0.04762
9	133	140	2.105	11.543	11.366	0.01112	3.702	4.172	0.01482
10	124	138	2.226	8.053	7.070	0.00869	3.010	3.540	0.02914
11	31	30	1.935	4.353	4.357	0.00000	1.590	1.969	0.07475

Table 3.1: Low voltage samples from the Northern Netherlands power grid compared with Random graphs of the same size.

- The edges are considered undirected.

These assumptions are common in power grid analysis from a graph theoretic perspective, see for instance [56, 185, 223, 222, 186, 57, 89] and lead to the following definition.

3.1. DEFINITION (POWER GRID GRAPH). *A* Power Grid graph *is a graph $G(V, E)$ such that each element $v_i \in V$ is either a substation, transformer, or consuming unit of a physical power grid. There is an edge $e_{i,j} = (v_i, v_j) \in E$ between two nodes if there is physical cable connecting directly the elements represented by v_i and v_j.*

The next step is to bring cable properties into the graph definition.

- For each cable connecting elements in the grid a weight is defined based on the multiplication of the following quantities:
  - The principal resistance characterizing the cable (whose value is given in Ohm/km).
  - The length of the cable (whose value is given in km).
- A special kind of connection is defined in the power grid known as a 'link'. These are connections, usually very short, with negligible resistance for which the specific value is not provided in the dataset. For edges representing these

links a conventional weight of 10^{-9} is given. This does not affect the overall validity of the weighted model since the number of links in a sample is extremely limited (about 1% of the overall connections are made of links).

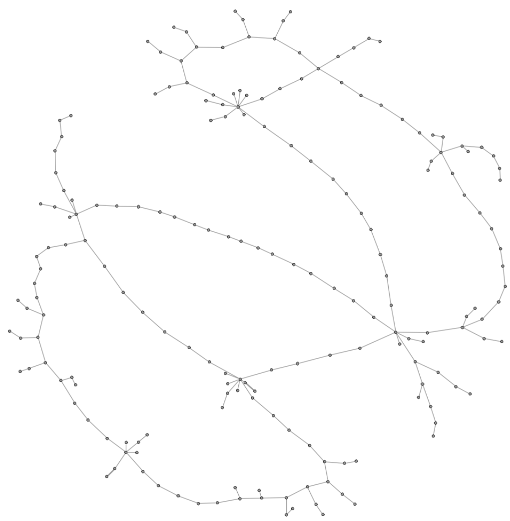
Two examples of graphs belonging to the low voltage and medium voltage network are shown in Figure 3.2.

3.2. DEFINITION (WEIGHTED POWER GRID GRAPH). *A* Weighted Power Grid graph *is a Power Grid graph $G_w(V, E)$ with an additional function $f : E \to \mathbb{R}$ associating a real number to an edge representing a physical property (e.g., the resistance, expressed in Ohm), of the physical cable represented by the edge.*

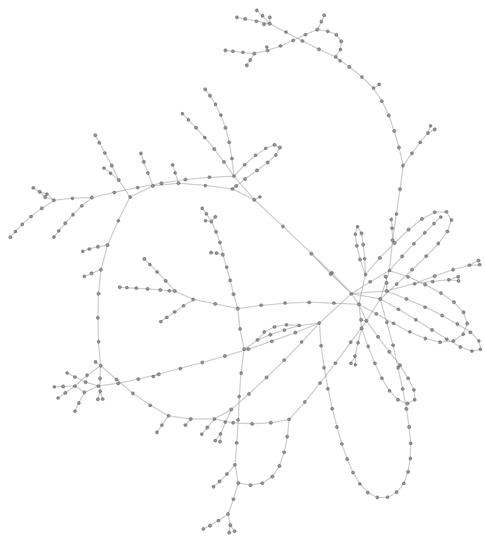
The analysis we perform uses samples from the low voltage and medium voltage grids. The low voltage samples sum up to a total of 663 nodes and a 683 edges; while the medium voltage samples sum up to 4185 nodes and a 4574 edges. The size of the data set, tough being a sample and not the whole network, is about the same size or larger than those used in other available studies on the (high voltage) power grid [223, 197, 57, 89, 184, 185, 186]. We begin our analysis by considering the unweighted model to derive basic topological properties and then proceed with a richer investigation by introducing graph weights.

	PRESENT STUDY						RANDOM GRAPH		
ID	Order	Size	Avg. d	APL	CPL	γ	APL	CPL	γ
1	444	486	2.189	11.033	10.858	0.00537	5.547	6.163	0.00333
2	472	506	2.144	17.095	17.174	0.01360	5.039	5.700	0.00106
3	238	245	2.059	11.715	11.580	0.00000	3.558	4.234	0.00595
4	263	288	2.190	12.775	12.311	0.01118	5.046	5.368	0.01080
5	217	229	2.111	10.321	10.241	0.00140	4.894	5.391	0.00121
6	191	207	2.168	9.288	8.990	0.00296	4.616	5.079	0.00225
7	884	1059	2.396	9.817	9.527	0.00494	5.440	6.010	0.00170
8	366	382	2.087	15.113	14.546	0.00000	4.691	5.249	0.00405
9	218	232	2.128	10.850	10.915	0.00000	5.454	5.856	0.00539
10	201	204	2.030	15.742	15.257	0.00166	4.898	5.503	0.00491
11	202	213	2.109	13.504	12.891	0.00140	4.801	5.217	0.08750
12	464	499	2.151	13.144	12.703	0.00036	4.718	5.390	0.00209

Table 3.2: Medium voltage samples from the Northern Netherlands power grid compared with Random graphs of the same size.



(a) Physical sample low voltage #5.



(b) Physical sample medium voltage #8.

Figure 3.2: Distribution grid samples.

3.2 Unweighted Power Grid Study

The typical study of the power grid as a complex system considers high voltage samples for characterizing the network structural properties and identifying how fragile the infrastructure is. We use similar techniques for the medium and low voltage. Let us begin by recalling the basic complex network quantities.

Basic properties and small-world analysis

We consider the traditional measures and metrics used in the literature of CNA, described in Appendix A and apply them to the data of the Dutch distribution grid. We divide our data set in samples of topologically connected regions. In Table 3.1, we report the analysis on the network data modeled as unweighted graphs and we compare each sample belonging to the low voltage network with a Random Graph of the same *size* and *order*. The analysis for the medium voltage is reported in Table 3.2. Referring to the table, the first column is the ID of the sample, the second and third represent the number of vertexes N (*order*) and edges M (*size*), respectively. The average degree (fourth column) is defined as $\langle k \rangle = \frac{2M}{N}$. The fifth and sixth columns report the average path length (APL) and characteristic path length (CPL), that is the average of the minimum distance between any two given nodes and the median of the average minimum distance between all node pairs, respectively. For more detailed definitions we refer to Definitions A.12 and A.13 on page 161. The seventh column provides an indication of the clustering coefficient (CC) of the graph, that is, broadly speaking, an average value of the ability of nodes to participate in connected aggregation with other nodes close to it. For a more detailed definition we refer to Definition A.8 on page 160. In general for more formal definitions related to graph theory concepts we refer to Appendix A.

We remark that the average node degree has similar values for the samples in the low and medium voltage, the values are both around 2. Computing the mean over all samples' average node degree gives a value of $\overline{\langle k \rangle} = 2.074$ with a very small variance $\sigma_{\langle k \rangle} = 0.017$. This value appears to be almost constant considering the low voltage and medium voltage samples since the variance of the two categories is even smaller ($\sigma_{\langle k \rangle_{LV}} = 0.016$, $\sigma_{\langle k \rangle_{MV}} = 0.012$). An almost constant average degree is also characteristic of the high voltage power grid [185], though with a slightly higher value $\langle k \rangle \cong 2.8$. This limited number of edges a node can manage can be regarded as a physical limit that each power grid substation has to satisfy.

Considering path measures: average path length and characteristic path length of the low voltage segment of the network have generally a smaller path length compared to the medium voltage one. The clustering coefficient is very small especially for the low voltage network for which many samples have a zero value (i.e., absence of triangles in the graph). The difference in path length between the low voltage network and medium voltage network is due to the higher number of nodes the medium voltage network samples have while holding the same average node degree as the low voltage, together with the absence of long distance edges. This implies a longer path to connect any two nodes in a bigger network. In addition, these values of APL and CPL are in general quite high, if compared to other networks such as the World Wide Web.

The clustering coefficients for the low voltage segment of the network are generally small; this is due to the strong hierarchical design of this layer of the physical network which resembles a tree-like radial structure. Contrarily, the medium voltage segment generally presents higher values for the clustering coefficient. This finding can be justified by the different purpose the medium voltage network has in which meshed components and connection redundancies are much more likely to be present for robustness reasons.

To gain a better understanding of the tables just presented, it is useful to compare the numbers obtained with those of Random Graphs [67] and to identify the possible presence of *small-world* properties. Small-world networks (SW), proposed by Watts and Strogatz in [223], own two important aspects at the same time: characteristic path length close in value to the one of a Random Graph (RG) ($CPL_{SW} \approx CPL_{RG}$) and a much higher clustering coefficient ($CC_{SW} \gg CC_{RG}$). Small-worlds are a better model than random graphs for social networks and other phenomena, and thus a candidate for modeling the power grid too. To make the comparison fair, random graphs are generated with the same number of nodes and edges as the real samples, imposing the resulting graphs not to have disconnected components. The values are presented on columns eight to ten of Tables 3.1 and 3.2. We note how the CPL of the grid samples is on average twice as big as the random generated samples, thus comparable to the definition of small-world graph according to [223]. In addition the clustering coefficient of the grid samples is almost always smaller than the result obtained for the random generated samples; this completely contradicts the definition of small-world graph according to [223]. Watts and Strogatz [223] impose the following condition to the graphs they study: $N \gg k \gg ln(N) \gg 1$ where N is the number of nodes, k is the number of edges per node. Such a condition is not satisfied by the Northern Netherlands samples and generally it is not satisfied by power grid networks as pointed by Wang *et al.*in [218]. Interestingly, the same condition is also not satisfied by the Western States high voltage power grid Watts and Strogatz use in [223] and Watts analyzes in [222], while the results for CC and CPL satisfy the conditions for a small-world network. Another study (i.e., [185]) considering the European high voltage power grid shows that the small-world phenomenon is not shown by all the considered grids, since especially the smaller (in terms of *order* and *size*) grids fail to satisfy the clustering coefficient condition.

In summary, the Northern Netherlands medium voltage and low voltage samples show a very small value of average node degree. This is mainly independent from the *order* and the different purpose of the network, being almost constant despite the different samples considered. In addition, the path length is quite high, given the *order* of the graphs, compared with other types of complex networks e.g., the World Wide Web. This relative high path length together with very small clustering properties suggests that the networks analyzed do not strictly follow the definition of small-world or, in terms of decentralized energy negotiation, it suggests that perhaps a structural change to decrease path length (especially the weighted one) might be necessary to empower delocalization.

Node degree distribution analysis

To have a general understanding of the overall characteristics of a network it is useful to compute certain statistical measures, one of which is the node degree probability distribution. The shape of the distribution is a salient characteristic of the network. For the power grid, the shape is typically either exponential or a power-law. More precisely an exponential node degree (k) distribution has a fast decay in the probability of having nodes with relative high node degree and follows a relation:

$$P(k) = \alpha e^{\beta k} \tag{3.1}$$

where α and β are parameters of the specific network considered. While a power-law distribution has a slower decay with higher probability of having nodes with high node degree:

$$P(k) = \alpha k^{-\gamma} \tag{3.2}$$

where α and γ are parameters of the specific network considered.

Power-law distributions are very common in many real life networks both created by natural processes (e.g., food-webs, protein interactions) and by artificial ones (e.g., airline travel routes, Internet routing, telephone call graphs), [13]. Having a power-law distribution in the node degree statistical distribution means that few nodes have a very high degree and the majority of nodes have very small degree. The types of networks that follow this property are referred as Scale-free networks ([16, 14, 5]); typical examples of Scale-free networks are the World Wide Web, the Internet, metabolic networks, airline routes and many others. From the dynamic point of view, the model that enables to create such networks is a preferential attachment model, that is, when new nodes and arcs are added to a graph, these are more likely to connect to nodes that already have a high degree, [15, 3]. In addition, this network structure provides special reliability properties: high degree of tolerance to random failures and, at the same time, very high sensitivity to targeted attacks towards hubs [5, 138, 56].

We compute the node degree distribution for every sample both for low voltage and medium voltage segments. For the most significant samples i.e., those belonging to the medium voltage and the big ones belonging to the low voltage part, the node degree cumulative distribution seems to follow, at least in a part of the distribution, a power-law: $P_k \sim k^{-\gamma}$. A coarse, but straightforward, method to investigate if the node degree follows a power-law is to plot the cumulative node degree distribution on a log-log scale [6]. If the distribution in a log-log plot follows a straight line the distribution can be considered a power-law, while if the decay is faster this might indicate an exponential distribution. It is also possible to apply data fitting techniques (e.g., maximum likelihood estimator and Kolmogorov-Smirnov test) to identify the γ parameter of a power-law as suggested by Clauset *et al.* [47].

The most significant samples for this kind of analysis are the biggest samples belonging to the medium voltage network and the most numerous ones from the low voltage (i.e., samples #5, #9 and #10 from Table 3.1). All these samples tend to follow a straight line in the log-log plot as in Figure 3.3. In the figure circles represent sample data, while the straight line represents a power-law with $\gamma = 1.977$. Power grids are not dense network and also many of the samples under investigation have small *order*, thus providing few data points. Therefore, for many samples it is difficult to perfectly fit to a distribution for the node degree following the usual node degree definition.

We thus conclude that only some samples of the medium voltage and low voltage tend to be Scale-free networks, although some exponential tail appear due to physical and economic constraints in the network. This means robustness in terms of redundancy of paths, but fragility to attacks on the hubs. The hubs tend to be the few nodes that most likely lead to the high voltage segment in a certain geographical location.

3.2.1 Betweenness analysis

Betweenness describes the importance of a node with respect to minimal paths in the graph. This property is important to identify critical components of the power grid [5, 138, 57]. For a given node, betweenness, sometimes also referred as *load*, is defined as the number of shortest paths that traverse that node. For a more technical definition refer to Definition A.15 on page 162. The betweenness measure allows to find if there are nodes that are critical for the whole infrastructure. In fact, the removal of nodes with the highest betweenness can lead to critical effects on the

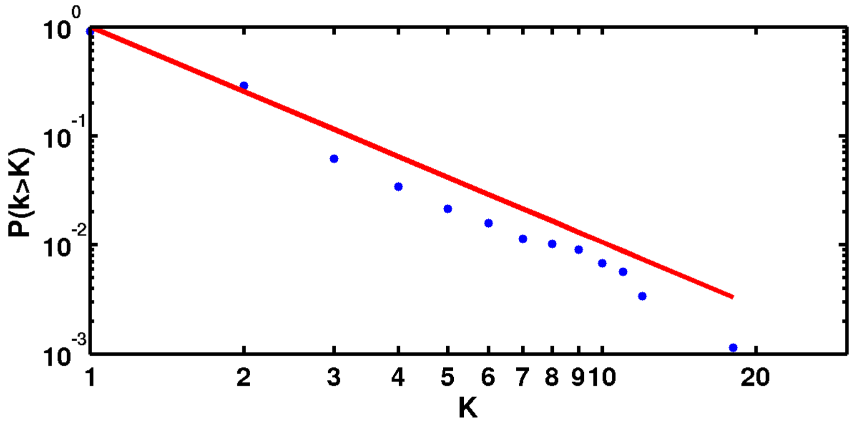


Figure 3.3: Node degree cumulative probability distribution for medium voltage sample #7.

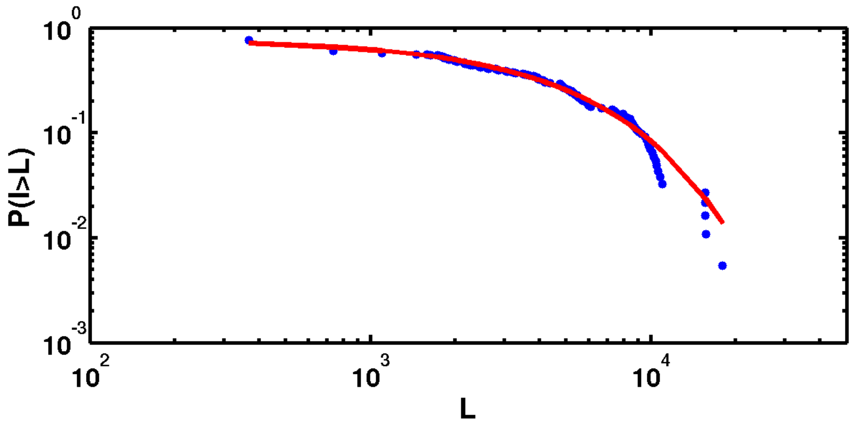


Figure 3.4: Betweenness cumulative probability distribution for low voltage sample #5.

network connectivity [138].

For the most significant samples in the low voltage network (i.e., samples #5 and #10) betweenness probability distribution follows an exponential decay, that is, the nodes with very high values of betweenness are less likely to be present in the network, as shown in Figure 3.4. In the figure circles represent sample data, while the continuous line represents an exponential decay $y = 0.7699e^{-2.227 \cdot 10^{-4}x}$. This aspect is not surprising since the low voltage network is quite hierarchical and the paths tend to follow the few ones admissible by the relative simple topology. In fact, the

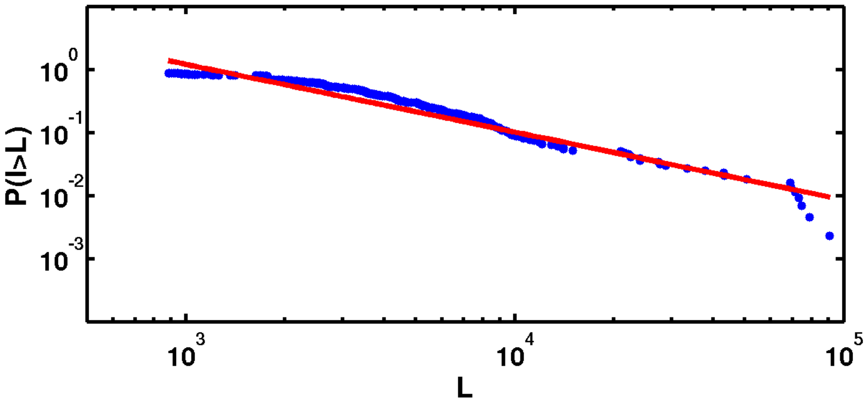


Figure 3.5: Betweenness cumulative probability distribution for medium voltage sample #1.

betweenness probability distribution charts (Figure 3.4) do not fit a straight line in a logarithmic plot, thus exhibiting a fast decay. In addition a fitting procedure, using the non-linear least square method gives very good results approximating the betweenness probability distribution samples with an exponential function. On the other hand, betweenness in medium voltage segment seems to follow a power-law decay, this is shown in the logarithmic chart in Figure 3.5. In the figure circles represent sample data, while the straight line represents a power-law with $\gamma = 1.075$. The samples from medium voltage network show a distribution of betweenness with a much fatter tail than the low voltage ones, that is there are several nodes that are central in many paths. This is due to the more meshed structure the medium voltage network has, compared to the low voltage one. This result for medium voltage network betweenness is closer to the results obtained for this same metric in high voltage studies, [5, 57]. In summary, a few nodes are extremely critical to enable the electricity distribution for the whole network.

3.2.2 Fault tolerance

A way to study the reliability of the network is by analyzing its connectivity when nodes are removed. There are basically two ways to perform this analysis: choosing the nodes randomly or selecting the nodes to be removed following a specific significant property for the network. Similar studies concerning the resilience of the high voltage power grid exist, e.g., [185, 5]. We apply such technique to the medium voltage and low voltage ends of the power grid using three policies for node removal:

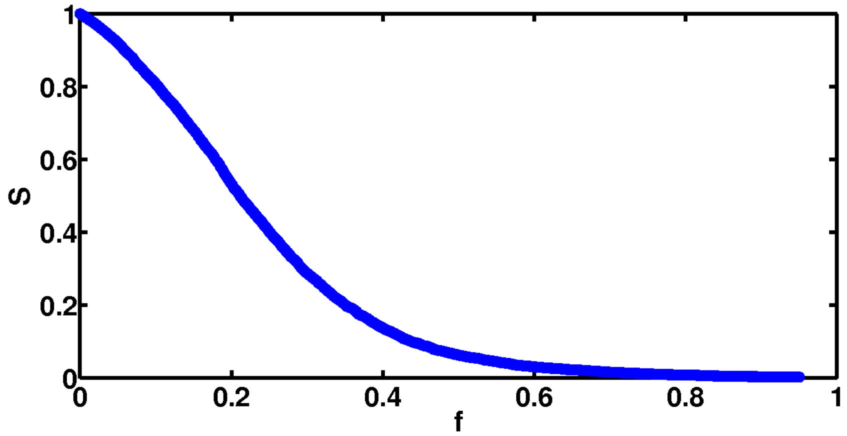


Figure 3.6: Resilience for node random-based removal for medium voltage sample #7.

random, highest degree and highest betweenness driven choices. The measure that is taken into account is the *order* of the largest connected component of the network (i.e., the number of nodes composing the biggest cluster in the network) computed as a fraction of the original *order* of the network, and its evolution while nodes of the network are removed, again the latter are considered as a fraction of the original *order* of the network.

The *random removal* simulates casual errors. As shown in [48], networks that follow a power-law whose characteristic parameter $\gamma < 3$ tend to have a high value for the transition threshold at which they disrupt. In the samples here analyzed, it seems that this is true especially for the small samples that generally have a cluster that is 10% of the original when almost 90% of the nodes are removed. The situation is different for samples with higher *order* that show a cluster that is reduced to 10% of the original when about 40% of the nodes are removed, as shown in Figure 3.6. Even if the degree distributions found for samples following a power-law have a parameter $\gamma < 3$ the samples show a threshold effect that is more similar, according to [48], to networks whose characteristic $\gamma > 3$.

The situation is radically different when "targeted attacks" are considered. In particular two kind of attack policies are investigated: *node degree-based removal* and *betweenness-based removal.* The main difference compared to the random-based removal is the presence of very sharp falls that appear when certain nodes are targeted. The removal of selected nodes can cause a drop in the size of the maximal connected component even of 40%. Node degree-based removal is much more crit-

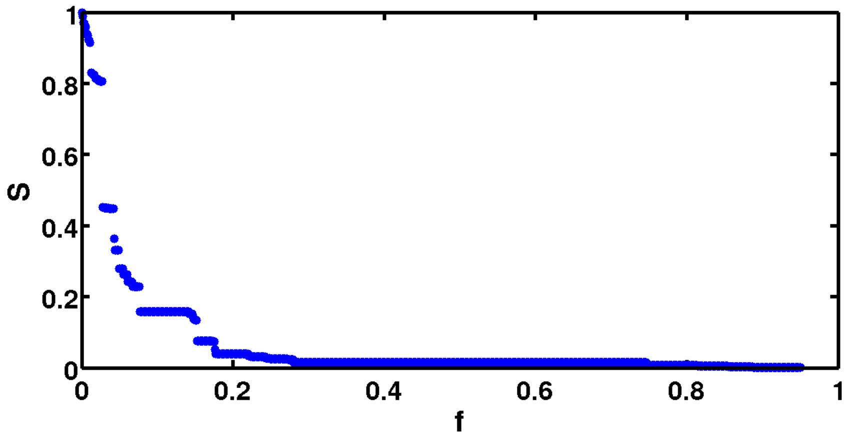


Figure 3.7: Resilience for node degree-based removal for medium voltage sample #7.

ical than the random removal: by just removing 10% of the most connected nodes one reduces the network to only 10% of its original size. The same applies for the biggest samples considered both in the low voltage and medium voltage network, as shown in Figure 3.7.

The removal of nodes based on the highest betweenness shows generally the same behavior, as the degree-based removal, with network disruption that appears much faster than random-based network failures. Considering the general correlation between nodes with a certain degree and their betweenness, it is not surprising that the two removal policies have very similar results and shape. The only remark that generally differentiates the betweenness-based removal is a little higher order of the maximal connected component compared to the one obtained with a degree-based removal when the same fraction of nodes is removed. In addition, the decrease of the order of the maximal connected component tends to be slightly smoother than the degree-based one. Figure 3.8 shows the comparison of the two removal policies for the samples that show some interesting deviations in the correlation of the degree and betweenness.

In summary, the results for the low voltage and medium voltage show disruption behaviors that are quite immune to random failures to which the networks present a constant degrading disruption, while they strongly suffer from the removal of specific nodes.

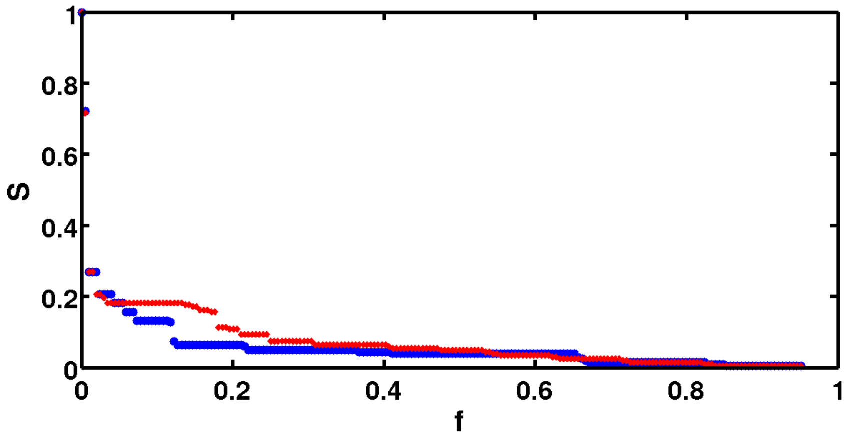


Figure 3.8: Resilience for node degree-based and betweenness-based removal for medium voltage sample #10.

3.3 Weighted Power Grid Study

The purely topological study of the power grid just presented already gives important information about the connectivity and robustness of the medium and low voltage grids, though it does not consider the different physical properties of the cables. These can vary greatly for different sections of the grid and provide essential indications to establish the behavior of a link. Next, we perform an analysis of the same samples of the grid considering the resistance as the weight of the graph model.

We take the samples analyzed in Tables 3.1 and 3.2, but now consider the weighted graph definition. The notion of a characteristic path length can be extended to take the weights into account yielding the values shown in Tables 3.3 and 3.4. In each table, the second column contains the characteristic path length resulting in the weighted graph (WCPL), cf. Definition A.14. The third column provides the average value of the weights of all edges; while the fourth column shows a normalized value for the weighted characteristic path length (NWCPL) obtained by dividing the WCPL by the average weight of the edge belonging to the same data sample.

Due to the relative short length of the low voltage networks cables, the WC-PLs for this segment of the network are small, as well as the average weight of each edge (almost all of them are below the unit). The situation is different for the medium voltage networks that have higher values since the cables and paths span across wider geographical areas. The discrepancy can be explained by the different

ID	Weighted Characteristic Path Length	Edge Average Weight	Normalized Weighted Characteristic Path Length
1	2.000	0.698	2.865
2	1.429	0.595	2.402
3	3.066	0.739	4.149
4	3.087	0.699	4.414
5	12.136	0.741	16.378
6	3.889	1.648	2.360
7	4.162	0.348	11.960
8	5.112	0.876	5.836
9	7.872	0.583	13.503
10	6.407	0.785	8.162
11	2.967	0.592	5.012

Table 3.3: Weighted analysis of the low voltage samples.

ID	Weighted Characteristic Path Length	Edge Average Weight	Normalized Weighted Characteristic Path Length
1	153.402	8.608	17.821
2	163.067	9.217	17.692
3	127.258	7.122	17.868
4	134.661	13.106	10.275
5	187.084	16.382	11.420
6	185.916	12.779	14.549
7	108.011	11.851	9.987
8	148.058	7.193	20.584
9	99.385	7.421	13.392
10	126.845	6.850	18.518
11	92.060	8.764	10.504
12	38.084	6.915	5.507
13	232.475	13.810	16.834

Table 3.4: Weighted analysis of the medium voltage samples.

purpose for which these networks are designed: a bridge network from high voltage transmission lines and end user distribution (medium voltage network) and the final end delivery (low voltage network). In fact, both the WCPL and the edge average weight for medium voltage samples are approximately two order of magnitude greater than the low voltage ones. This is indeed due to an extension of medium voltage cables that range from hundred meters to kilometers, while low voltage extend usually around tens of meters.

Weighted node degree distribution

In the weighted power grid graph that we consider, there are no weights explicitly associated to the nodes. However, the weights of the incident edges influence the node properties. To capture this characteristic one can consider a weighted node degree for a node whose weight is obtained by the sum of the weights of the incident edges. For a more formal definition we refer to Definition A.7 on page 160. It is then easily to extend the concept of node degree distribution to a weighted node degree distribution. This distribution is straightforwardly obtained by using the weighted degree concept instead of the traditional node degree in the node degree distribution (cf. Definition A.16 on page 162).

For the most significant sample of the low voltage, as shown in Figure 3.9, the shape of the distribution is close to an exponential one with a quite fast decay. In the figure circles represent sample data, while continuous line represents a sum of exponential decays $y = 0.8975e^{-0.9289x} + 0.0904e^{-0.1379x}$. The situation looks different in medium voltage samples. The very first part of the distribution is well fitted by an exponential shape, while the central part of the distribution, and especially the tail, fit best a power-law like shape as visible in Figure 3.10. In the figure circles represent sample data, while straight line represents a power-law with $\gamma = 1.374$. An explanation of such behavior among the most numerous samples of the two ends of the grid is due to the *order* and *size* of the medium voltage samples that are from two to four times bigger than the low voltage samples, thus having a higher likelihood of far different values in weighted node degree.

3.3.1 Betweenness analysis

Considering the weighted definition of path it is possible to compute betweenness in the weighted scenario and again betweenness of a node can be seen as a random variable thus obtaining the corresponding probability distribution. The shape of the distribution does not change much compared to the same unweighted samples for the low voltage network: the distribution is best approximated by an exponential decay or by a sum of exponential contributions. For medium voltage samples, the changes between unweighted and weighted paths influence the betweenness probability distribution whose shape in these conditions seems to be better approximated by an exponential or sum of exponential components. This change in the distribution of the number of shortest paths that traverse a node between the weighted and the unweighted graph is clearly an indication that some property change between

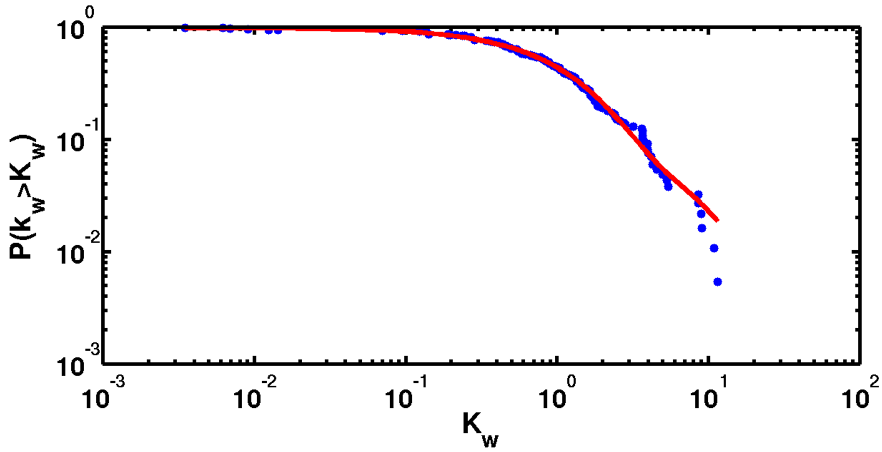


Figure 3.9: Weighted node degree cumulative probability distribution for low voltage sample #5.

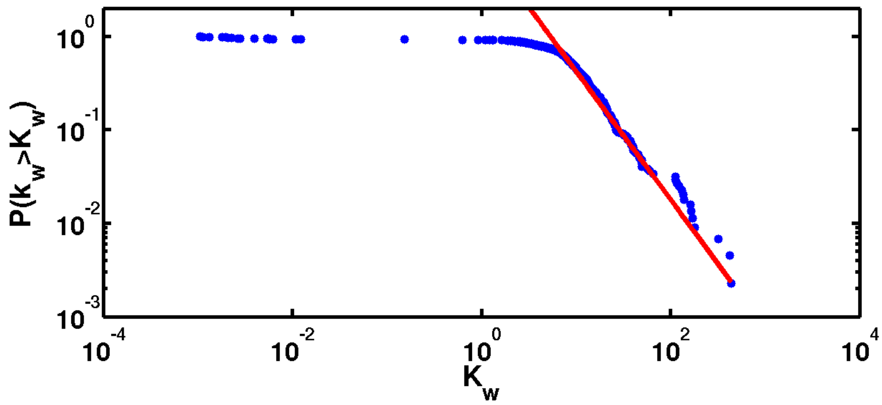


Figure 3.10: Weighted node degree cumulative probability distribution for medium voltage sample #1.

the two analysis. In fact, the weighted path analysis better approximates the actual routes the current flows follow.

3.3.2 Fault tolerance

Fault tolerance can be evaluated based on the removal of nodes following strategies similar to the unweighted case. Since the random removal yields exactly the same

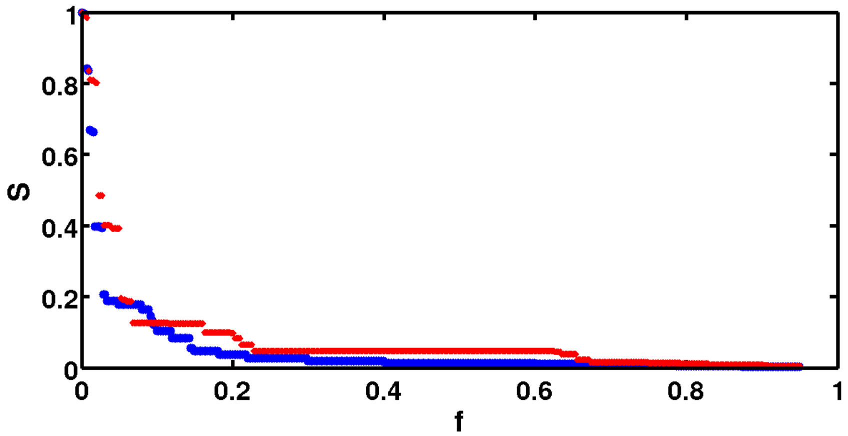


Figure 3.11: Resilience for node degree-based removal for medium voltage sample #1.

result for the weighted and unweighted case, here we focus on the node degree-based removal that considers the weighted node degree definition. The disruption behavior of the network samples is very similar to the unweighted situation: the network suffers deeply these targeted attacks; a very small percentage of removed nodes causes an important loss in the size of the biggest component left in the network. There is a general correlation between high degree nodes in the unweighted graph and high degree nodes in the weighted one. If one takes a closer look at the disruption charts for the same samples, some small differences can anyway be noted as in Figure 3.11. The horizontal axis represents the fraction f of the nodes removed from the original sample; the vertical axis represents the size of the largest connected component S relative to the initial size of the graph. Red diamonds represent the weighted node degree-based removal, while blue circles represent *traditional* node degree-based removal. The nodes with the highest weighted degree cause a bigger damage to the network when removed in the very first iteration than nodes with higher degree in unweighted networks. The situation then changes in the later stages of the removal process when a bigger disruption is caused by nodes with higher node degree in *traditional sense*.

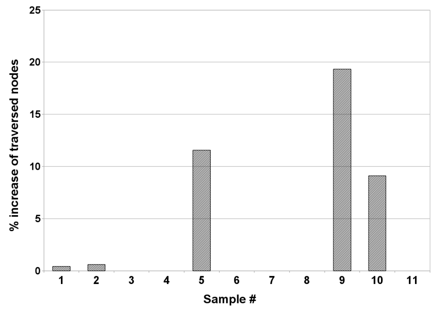
3.3.3 Unweighted vs. weighted study

In Chapter 2 and more thoroughly in [167, 87], the superiority of an approach that considers the physical properties underlying the network is emphasized. In this section and in Section 3.2 we have performed the same type of analysis considering both the unweighted and weighted definition of graphs, respectively. Our motivation was not to prove the superiority of one approach over the other, but to consider if important differences in the metrics analyzed were present.

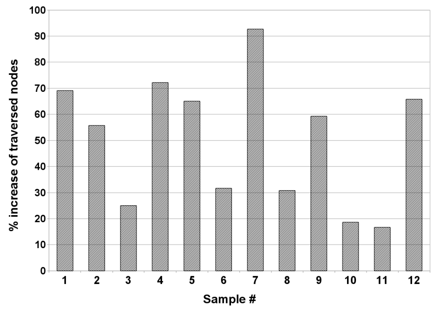
The similarities lie in the fashion the network disrupts when the most connected nodes (either in the traditional, or in the weighted definition) are attacked: the removal of just few nodes compromise the connectivity of the whole network (cf. Figure 3.11). Some dissimilarities are evident considering the node degree distribution. In fact, the weighted analysis tends to reduce the contribution of the tail components of the distribution, thus being more compact especially for low voltage samples. This is due to the small variance of the weighted node degree that especially these low-end samples of the grid show. Samples that in the unweighted analysis show a power-law distribution, when considered weighted tend to assume an exponential form or a sum of exponential contributions. The most interesting dissimilarity in the analysis is in the path-related properties. The number of nodes traversed in the two cases is different for the medium voltage networks. This is particularly interesting from the practical point of view, as it indicates the number of transformers and distribution substations traversed in the power grid. These points are critical in terms of additional losses that are associated with substations and transformers, and in turn in the number of potential points of failure that a path traverses. Figure 3.12 shows the results for the low voltage and medium voltage networks. Each bar represents the average percentage increase in the number of nodes traversed along the shortest path between any two nodes for the unweighted and the weighted situation. It is interesting that for several samples of the low voltage network there is no difference in the number of traversed nodes, thus reinforcing the idea of a highly hierarchical tree-like radial structure whose paths are fixed by the built-in topology of the grid independently of the associated edge weights. The situation though is quite different for the medium voltage. In fact, there is an increment of traversed nodes between the weighted and unweighted models (especially for the meaningful samples) on average of about 50%. This is a clear indication of a meshed network for which there are less imposed paths and in which weights have an important role.

3.4 Relating Grid Topology to Electricity Distribution Costs

The complex network analysis that we provided so far gives a statistical aggregated view of the current infrastructure for the low and medium voltage grids. The natural next question that arises concerns the usability of such infrastructure for the



(a) Low voltage samples



(b) Medium voltage samples

Figure 3.12: Percentage increase in number of node traversed between weighted and unweighted graph definition.

delocalized energy exchange. To answer it, we propose to tie statistical properties of the power grid with energy trading costs. The cost represents a balance frontier below which the actors are motivated to trade energy. We do not claim of having identified "the" cost function, but rather we propose that complex network analysis measures can be used to evaluate the success of a decentralized energy market.

In general, establishing energy pricing is not a simple task since several aspects influence the overall price at which electricity is sold [84]. There are aspects connected to the supply side such as fuel prices, policy regulations, load losses and bidding strategies; on the other hand, there are elements connected to the demand side such as human behaviors, natural phenomena that influence habits and thus consumption. Recent proposals and methods for price allocation include *nodal pricing* [119], which is particularly indicated for distributed generation solutions because of the price benefits it brings to the small producer [194]. It is also important to notice that the savings deriving from distribution losses can be extremely important [43, 104, 115]. A set of factors is most definitely tied to infrastructural properties of the distribution network, as illustrated for instance in the economic studies of Harris and Munasinghe [84, 141], most notably:

- losses both in line and at transformer stations,
- security and capacity factors,
- line redundancy, and
- power transfer limits.

The listed technical parameters are naturally associated with a topological parameter, namely:

- Line losses are related to and thus expressed as a function of the weighted characteristic path length $L_{line_N} = f(WCPL_N)$
- Substation losses are expressed as a function of the (average) number of nodes visited while traveling from source node to destination node along the weighted shortest path $L_{substation_{ij}} = f(|WSP_{ij}|)$. The significant dimension is $L_{substation_N} = f(\overline{|WSP_N|})$ where $\overline{|WSP_N|}$ is the average number of nodes traversed in a shortest path in the network N.
- Robustness from a topological perspective is expressed as the fraction of the maximal connected component compared to its original *order* once a certain fraction of nodes is removed. $Rob_N = f(N, p)$ where N is the network under evaluation and p is the removal policy adopted.
- Line redundancy is simply mapped to a topological metric that counts the number of paths (without cycles) that are available between any two nodes and the cost associated to this redundancy $Red_{ij} = f(|P_{ij}|, w_{ij})$ where P_{ij} is the set of paths between nodes i and j and w_{ij} is the weight of the worst case redundant path. A global metric for network N is $Red_N = f(\overline{|P_{ij}|}, \overline{w_{ij}}), \forall i, j \in N$
- Network capacity may be considered as a function of the weighted characteristic path length where the weight in this case is the maximal supported operating current of the cable $Cap_N = f(WCPL_N)$

These topological ingredients provide two sorts of measures, the first one α gives an average of the dissipation in the transmission between two nodes

$$\alpha = f(L_{line_N}, L_{substation_N}); \tag{3.3}$$

the second one β is a measure of reliability/redundancy in the paths among any two nodes

$$\beta = f(Rob_N, Red_N, Cap_N). \tag{3.4}$$

We argue that these two factors influence the inclination of prosumers (energy consumers/producers) to trade energy on the power grid. In fact a high value of the α parameter represents a high level of losses experienced for transporting energy in the network, either in distribution lines or substations. Additionally, the reliability and ability to bring sufficient energy to the end users plays an important role. In

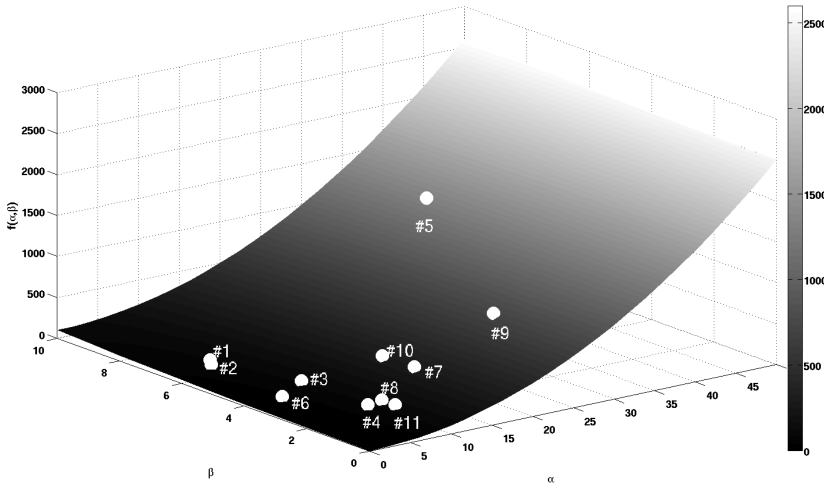


Figure 3.13: Transport cost of energy based on the low voltage topological properties.

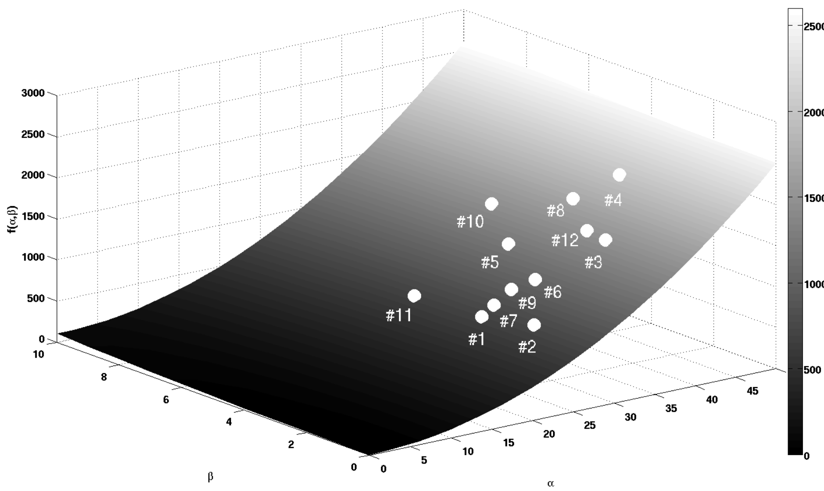


Figure 3.14: Transport cost of energy based on the medium voltage topological properties.

fact, if proper levels of robustness of the network or resilience to failures are not the norm, the prosumer inclination to sell energy as well as the end user to buy it will be limited. Furthermore, if the availability of redundant paths for electricity routing in case of partial disruption of the network is insufficient, then it leads to a high value of β, and in turn a disincentive for trading. To better understand the constituents of α and β, we consider next a possible instantiation of these parameters using the data of the Dutch power grid.

3.4.1 Application to the Dutch distribution grid

To give an impression of how the parameters can be used to assess the success of the energy market, we provide an example next.

- Losses on the transmission/distribution line can be expressed by the quotient of the weighted characteristic path length and the average weight of a line (a weighted edge in the graph):

$$L_{line_N} = \frac{WCPL_N}{\overline{w}} \tag{3.5}$$

- Losses at substation level are expressed as the number of nodes (on average) that are traversed when computing the weighted shortest path between all the nodes in the network:

$$L_{substation_N} = \overline{Nodes_{WCPL_N}} \tag{3.6}$$

- Robustness is evaluated with random removal strategy and the weighted node degree-based removal by computing the average of the *order* of maximal connected component between the two situations when the 20% of the nodes of the original graph are removed. It can be written as:

$$Rob_N = \frac{|MCC_{Random20\%}| + |MCC_{NodeDegree20\%}|}{2} \tag{3.7}$$

- Redundancy is evaluated by covering a random sample of the nodes in the network (40% of the nodes whose half represents source nodes and the other half represents destination nodes) and computing for each source and destination pair the first ten shortest paths of increasing length. If there are less than ten paths available, the worst case path between the two nodes is considered. To have a measure of how these resilient paths have an increment in transportation cost, a normalization with the weighted characteristic path length is performed. We formalize it as:

$$Red_N = \frac{\sum_{i \in Sources, j \in Sinks} SP_{w_{ij}}}{WCPL} \tag{3.8}$$

- Network capacity is considered as the value of the weighted characteristic path length, whose weights are the maximal operating current supported, normalized by the average weight of the edges in the network (average current supported by a line). That is:

$$Cap_N = \frac{WCPL_{currentN}}{\overline{w_{current}}} \tag{3.9}$$

With these instantiations, equations (3.3) and (3.4) become:

$$\alpha = f(L_{line_N}, L_{substation_N}) = L_{line_N} + L_{substation_N} \tag{3.10}$$

$$\beta = f(Red_N, Rob_N, Cap_N) = \frac{Red_N}{Rob_N \cdot ln(Cap_N)} \tag{3.11}$$

The functions to compute α and β are only few of many available possible ones. The choice made here is to have a simple mechanism to assess the potential distribution costs of different networks. Equation (3.10) is a basic sum, since both quantities are pure numbers, over the losses that are experienced both at line and at substation level. Equation (3.11) takes into account the aspects of reliability and tolerance of the network: the higher the β the more prone to failures and less reliable the network is. The cost of increasing the number of paths to provide more redundancy is the dividend in the fraction, while elements improving reliability act as divisor.

With these quantities, one can form an impression of what the influence of the cost of transportation is for the decentralized energy exchange. If the cost is too high because of an infrastructure with high chances of failure (high β) and high resistance (high α), then the decentralized market will not be incentivized. On the other hand, for low α and β, it will be economically attractive to have a decentralized energy market. In Figures 3.13 and 3.14, we show a combination of α, β obtained with the functions described above and with an hypothesis of a quadratic increase of energy price with the increment in α and β. We report the position of the analyzed samples as white circles in Figures 3.13 and 3.14 respectively for low voltage and medium voltage samples. By performing an economic study, which we stress is beyond the scope of the present treatment, one then can identify what the threshold is for the feasibility of a decentralized market (gray rectangles in Figure 3.15) and then conclude what topological modifications are necessary to the medium voltage and low voltage infrastructure in order to allow the energy exchange.

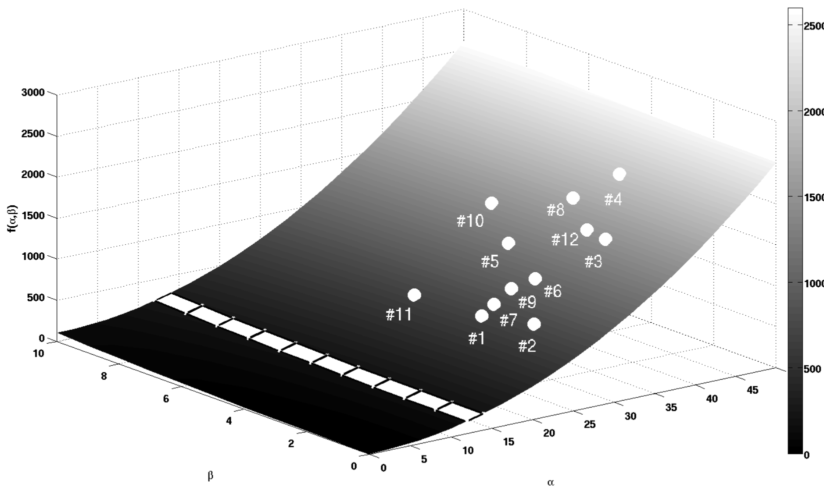


Figure 3.15: Transport cost of energy based on the topological properties for medium voltage with supposed economic convenience threshold (grey thick line).

Chapter 4

Network Evolutions for the Smart Grid

We have seen that complex network analysis (CNA) can provide a useful insight in understanding the topology of the distribution grid (cf. Chapter 3). In particular, we have used complex network analysis techniques not only to analyze the reliability of the network, as the majority of the literature studies concerning the high voltage grid, but we have also considered how the topology can play a role in influencing the price of electricity distribution.

Here we want to go one step further and consider possible scenarios of evolution of the distribution grid in terms of topology, since we believe that the future smart grid will have to evolve and a change in the distribution infrastructure. Inevitably, more generation and more electricity consumption at local-scale will have repercussions on the distribution grid infrastructure as well. More and more prosumers will feed their locally produced energy into the grid and participate in a local market for energy trade. The distribution grid has then to accommodate a new paradigm that is no more the top-down approach where the power is generated in big facilities remotely located, but it is dominated by a bottom-up approach where the producers are numerous and energy is produced and consumed at the neighborhood level. The distribution grid has therefore to be efficient in enabling this paradigm without posing barriers with high cost in energy transportation. In addition, the massive presence of electric vehicles will also be an important ingredient in shaping the next generation of distribution grids. In order to look at the evolutions that may characterize the future distribution grid, we consider here two situations to understand how the new grids could look like. First, we consider the case of the realization of completely new infrastructures such as for new settlements whose topology blueprint can be decided by an infrastructure planner. The blueprint to be used is based on well-known models of complex network analysis (Section 4.2). Second, we consider the case of current networks deployed on the ground and we look at techniques to adapt the distribution grid topologies to better support the local energy interactions of the smart grid (Section 4.3). The former can be considered as a solution for developing new distribution infrastructures in new locations e.g., new urbanization expansions in developing countries where the core is the local energy production and distribution, while the latter deals more with the changes required in the current distribution grid to become more efficient and better support energy exchanges that take place locally. The novelty is first in using the complex network analysis as a design tool for infrastructure planning and evolution, and, second, the focus that, compared to the existing studies of power grid and complex network analysis, is devoted to the distribution grid and not to the high voltage grid.

4.1 Network Metrics

The price of electricity distribution is influenced by the physical characteristic of the network that transports it. We have shown in Chapter 3 how different samples of the Dutch distribution grid have differences in their topology and the influence on the distribution price in that network. To measure those differences, we proposed a number of metrics useful for analyzing power grid topologies having in mind decentralized energy trading. Our goal has been to understand their influence on the electricity distribution price. Here we define a set of additional metrics inspired by the analysis carried out on the Dutch samples, and we then use them to measure properties of networks for the evolution of the distribution grid. We set two main categories of requirements: qualitative and quantitative desiderata that the networks optimized for the local energy exchange should satisfy.

Qualitative requirements

The main qualitative requirement we envision for the future distribution network concerns the modularity of the topology. In the power system domain, the modularity is invoked as a solution that provides benefits reducing uncertainties in energy demand forecasting and costs for energy generation plants as well as risks of technological and regulatory obsolescence [123, 88]. Modularity is usually required not only in the energy sector, but more generally in the design and creation of products or organizations [77]. It is also a principle that is promoted in innovation of complex systems [68] for the benefits it provides in terms of reduced design and development time, adaptation and recombination. We define the modularity as the ability of building the network using a self-similar recurrent approach and having a repetition of a form of pattern in its structure.

Quantitative requirements

As a global statistical tool, quantitative requirements give a precise indication of network properties. Here are the relevant ones when considering efficiency, resilience, and robustness of a power system.

- *Characteristic Path Length (CPL) lower or equal to the natural logarithm of order of graph:* $CPL \leqslant ln(N)$. This requirement represents having a general short path when moving from one node to another. In the grid this provides for a network with limited losses in the paths used to transfer energy from one node to another.

- *Clustering Coefficient (CC) which is 5 times higher than a corresponding Random Graph (RG) with same order and size*: $CC \geqslant 5 \times CC_{RG}$. Watts and Strogatz [223] show that small-world networks have clustering coefficient such that $CC \gg CC_{RG}$. Here we require a similar condition, although less strong by putting a constant value of 5. This requirement is proposed in order to guarantee a local clustering among nodes. In fact, it is more likely that energy exchanges occur at a very local-scale (e.g., neighborhood) when small-scale distributed energy resources are heavily implemented.

- *Betweenness-related requirements:*

  - *A low value for average betweenness in terms of order of the graph* $\overline{\upsilon} = \frac{\overline{\sigma}}{N}$, where $\overline{\sigma}$ is the average betweenness of the graph and N is the *order* of the graph. For the Internet, Vázquez *et al.* [210] have found for this metric $\overline{\upsilon} \approx 2.5$. Internet has proved successful to tolerate failures and attacks [48, 5], therefore we require a similar value for this metric for the future power grid.
  - *A coefficient of variation for betweenness* $c_v = \frac{s}{\overline{x}} < 1$ where s is the sample standard deviation and $\overline{x}$ is the sample mean of betweenness. Usually, distribution with $c_v < 1$ are known as low-variance ones.

  The above two requirements are generally considered to provide network resilience by limiting the number of critical nodes that have a high number of minimal paths traversing them. These properties provide distributions of shortest paths which are more uniform among all nodes.

- *An index for robustness such that* $Rob_N \geqslant 0.45$. Robustness is evaluated with a random removal strategy and a node degree-based removal strategy by computing the average of the *order* of the maximal connected component (MCC) of the graph between the two situations (i.e., random and targeted node removal) when the 20% of the nodes of the original graph are removed. It can be written as $Rob_N = \frac{|MCC_{Random20\%}|+|MCC_{NodeDegree20\%}|}{2}$. Such a requirement is about double the value observed for current medium voltage networks and 33% more for low voltage network samples.

- *A measure of the cost related to the redundancy of paths available in the network:* $APL_{10^{th}} \leqslant 2 \times CPL$. As described in Chapter 3, with this metric we consider the cost of having redundant paths available between nodes. In particular, we evaluate the 10^{th} shortest path (i.e., the shortest path when the nine best ones are not considered) by covering a random sample of the nodes in the network (40% of the nodes whose half represents source nodes and the other half represents destination nodes). The values for the paths considered are then averaged. In the case where there are less than ten paths available, the worst case path between the two nodes is considered. This last condition gives not completely significant values when applied to networks with small connectivity (i.e., absence of redundant paths).

Metric	**Efficiency**	**Resilience**	**Robustness**
CPL	✓		
CC	✓		
Avg. Betweenness		✓	
Betw. Coeff. of Variation		✓	
Rob_N			✓
$APL_{10^{th}}$	✓	✓	

Table 4.1: Metrics classification related to properties delivered to the network.

We categorize the above quantitative metrics into three macro categories with respect to how they affect the power grid and measure its goodness from a topological point of view: efficiency deals with the losses in the transfer of energy, resilience measures the possibility of having alternative paths if part of the network is compromised/congested, and robustness concerns failures happening to the overall network connectivity. Table 4.1 summarizes the property each metric assesses. Each metric gives a specific contribution and all the metrics together cover all the properties a smart grid infrastructure should have.

4.2 Building New Distribution Networks

To address the question of what are the best suited topologies to characterize the medium and low voltage grids, we study models for graph generation proposed for social and technological complex networks. For each model we evaluate the properties of the network for several values of the *order* of the graph. Following our analysis of the Northern Dutch medium and low voltage in Chapter 3, we categorize networks as *small*, *medium* and *large*, see Table 4.2. We then analyze the properties of the networks coming from the generated models by applying the relevant complex network analysis metrics described in Section 4.1. In this way, CNA is not only a tool for analysis, but it becomes a design tool for the future electrical grid.

Network layer	Category	*Order*	Reference Dutch sample
low voltage	Small	≈20	#1
low voltage	Medium	≈90	#7
low voltage	Large	≈200	#5
medium voltage	Small	≈250	#3
medium voltage	Medium	≈500	#12
medium voltage	Large	≈1000	#7

Table 4.2: Categories of medium and low voltage network and their *order*.

Synthetic models

We look at network models that have proven successful in showing salient characteristics of technological networks (i.e, preferential attachment, Copying Model, power-law networks), social networks (i.e., small-world, Kronecker graph, recursive matrix) and natural phenomena as well (e.g., Random Graph, small-world, Forest Fire) to investigate which one is best suited for supporting local-scale energy exchange from a topological point of view. Next, we provide a brief introduction to all the models used in the present study, while a more in-depth presentation is available in Appendix B or [39] and [151].

- **Random Graph.** A Random Graph is built by connecting each pair of nodes with an edge with probability p. It is due to the pioneering studies of Erdős and Rényi [67].
- **Small-world Graph.** The small-world characterization of graphs has been investigated by Watts and Strogatz [222, 223], who showed the presence of the small-world property in many types of networks such as actor acquaintances, the power grid and neural networks in worms.
- **Preferential Attachment.** The preferential attachment model represents the phenomenon happening in real networks where a fraction of nodes has a high connectivity while the majority of nodes has small node degree. This model is built upon the observation by Barabási and Albert [15] of a typical pattern characterizing several type of natural and artificial networks.
- **R-MAT.** R-MAT (Recursive MATrix) is a model that exploits the representation of a graph through its adjacency matrix [40]. In particular, it applies a recursive

method to create the adjacency matrix of a graph, thus obtaining a self-similar graph structure. This model captures the community-based pattern appearing in some real networks.

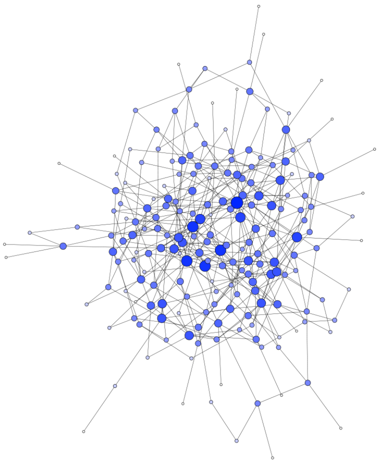
- **Random Graph with power-law.** A Random Graph with power-law model generates networks characterized by a power-law in the node degree probability distribution ($P(k) \sim k^{-\gamma}$). A graph generated in this fashion has the majority of nodes with a low degree and a small amount of nodes with a very high degree. Power-law distributions are very common in many real life networks both created by natural processes (e.g., food-webs, protein interactions) and by artificial ones (e.g., airline travel routes, Internet routing, telephone call graphs) [13].

- **Copying Model.** Replicating the structure underlying the links of WWW pages brought to the Copying Model [107] which captures the tendency of members of communities with same interests to create pages on the Web with a similar structure of links.

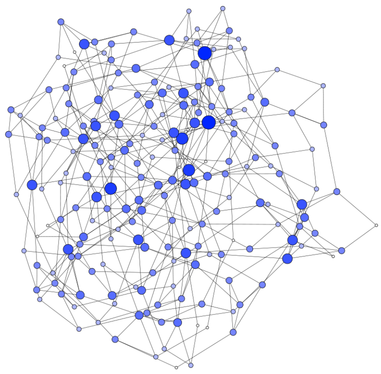
- **Forest Fire.** In order to capture dynamic aspects of the evolution of networks, Leskoveck *et al.* [118] proposed the Forest Fire model. The intuition is that networks tend to densify in connectivity and shrink in diameter (i.e., the greatest shortest path in the network) during the growth process. Several technological, social and information networks exhibit this property.

- **Kronecker Graph.** A generating model with a recursive flavor similar to R-MAT uses the Kronecker product applied to the adjacency matrix of a graph [117]. If the Kronecker product is applied to the same matrix, therefore multiplying the matrix with its elements recursively, a self-similar structure arises in the graph. This model creates networks that show a densification in the connectivity with a shrinking diameter over time.

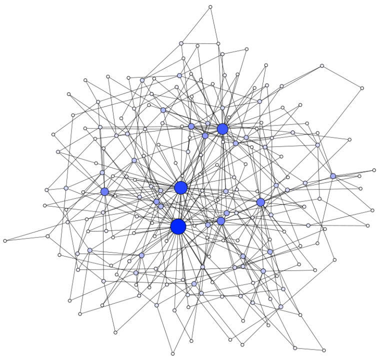
An example of graphical representation of the topological models obtained by applying the generation algorithm of each network model just described is shown in Figures 4.1 and 4.2.



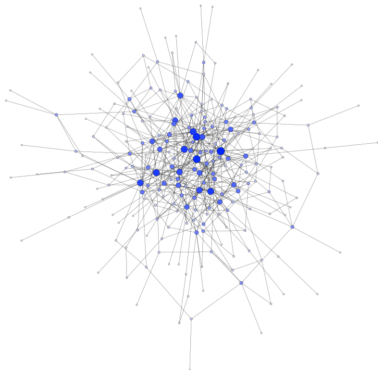
(a) A Random Graph (199 nodes, 400 edges).



(b) A small-world graph (200 nodes, 399 edges).

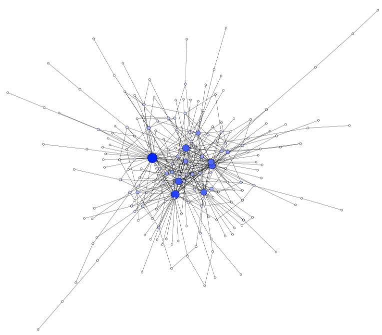


(c) A preferential attachment graph (200 nodes, 397 edges).

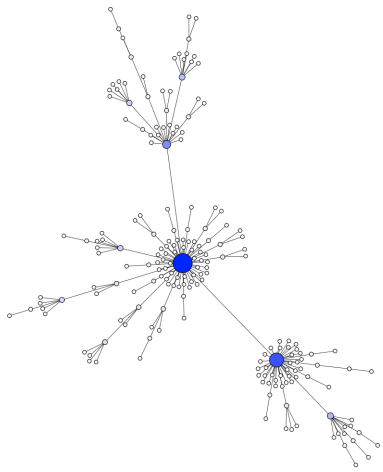


(d) A R-MAT graph (222 nodes, 499 edges).

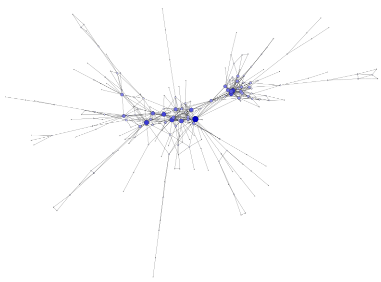
Figure 4.1: Graphical representation of the network models considered in the study.



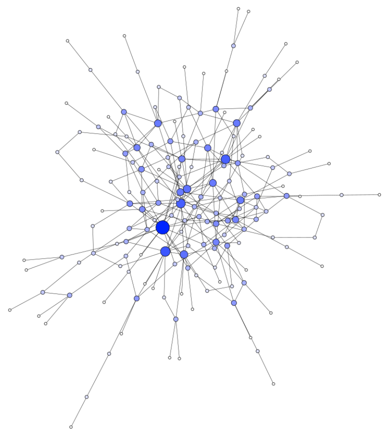
(a) A Random Graph with power-law graph (200 nodes, 399 edges).



(b) A Copying Model graph (200 nodes, 199 edges).



(c) A Forest Fire graph (200 nodes, 505 edges).



(d) A Kronecker graph (167 nodes, 264 edges).

Figure 4.2: Graphical representation of the network models considered in the study.

Model generation and metrics computation.

The synthetic topologies that we generate come from complex network analysis literature. These networks are obtained using the Stanford Network Analysis Project

(SNAP)[1] library that implements the generation algorithms of the network topologies described above and more thoroughly in Appendix B. The analysis of the generated graphs according to the metrics described in Section 4.1 is performed with ad hoc created software based on the JAVA graph library JGraphT.[2] The versions of SNAP and JGraphT software libraries used are respectively v10.10.01 and v0.8.1.

Analysis and metric satisfaction of current medium and low voltage networks

The baseline network for comparing possible evolutions must be the current power grid. Therefore, we use actual samples from the medium and low voltage network of the Northern Netherlands (cf. Chapter 3). The average degree of the medium and low voltage samples scores almost constantly around $\langle k \rangle \approx 2$ independently of the *order* of the network. In the low voltage networks we see a tendency towards the increase of the characteristic path length, with a value of about 18 when the *order* and *size* tend to 200 nodes and edges, respectively. The metric does not have the same clear tendency for the medium voltage samples. Considering the clustering coefficient there is a general rule: a null value for the low voltage samples and small, but at least significant, values for the medium voltage samples. These differences in both characteristic path length and clustering coefficient come from the difference in topology of the two networks. The low voltage network is almost a non-mashed network which resembles for certain samples trees, closed chains or radial structures with longer paths on average, especially for networks with bigger *order*. On the other side, the medium voltage network is more meshed (despite the same average node degree) with more connections that act as topological "shortcuts". It also has some redundancy in the connections between the neighborhood of a node, which implies a significant clustering coefficient compared to the low voltage network. The analysis of the robustness metric shows generally poor scores that decrease while the samples increase in *order*, at least for the low voltage networks, while the tendency is not clear for the medium voltage samples considered. A common behavior for the medium voltage samples is the problem they experience in the biggest component connectivity, when the 20% of the nodes with the highest degree are removed from the network the *order* of the MCC falls to 4.56%, 3.66% and 3.96% of its initial value respectively for the *small*, *medium* and *large* samples (cf. reference sample in Table 4.2). Considering the additional effort required when the first nine shortest paths are not available, we see a general increase especially for the low voltage samples, where the $APL_{10^{th}}$ increases three times for the *large* sample analyzed; the increase is still present in the medium voltage networks, but it is limited when

[1] `http://snap.stanford.edu/`
[2] `http://www.jgrapht.org/`

compared to the low voltage samples. It is again an indication that the medium voltage provides more efficient alternative paths to connect nodes. An exception in the results is the low voltage *medium* size sample: here the $APL_{10^{th}}$ is very close to the traditional characteristic path length. This is due to the absence of alternative paths, therefore the only paths between nodes are at the same time the best and worst case too. Such aspect reinforces the idea of a low voltage network with a fixed structure (sort of chain, tree-like or radial structures) and a limited redundancy.

Considering the betweenness-related metrics (Table 4 in [160]), one notes an increase in the average betweenness as the samples become more numerous in the two segments of the network. This same tendency is present in the average betweenness to *order* ratio: the biggest samples in terms of *order* both of low voltage and medium voltage score highest. In particular, the *large* sample belonging to the low voltage is almost twice the value of the biggest sample of the medium voltage. Again it can be justified by the similar-to-tree structure of the low voltage samples for which nodes responsible for the paths that enable sub-trees or sub-chains to be connected are the most high scoring for betweenness. This tendency highly increases the average betweenness (while the mode is usually null). The coefficient of variation is above one for all the big samples and reaches almost three for the biggest sample belonging to the medium voltage network. Such an high value implies an high standard deviation in the betweenness of the nodes.

4.2.1 Model parameters

To model the future power grid we compare network topologies that evolve in *size* (M) and *order* (N). In particular, we consider the increase of average node degree ($\langle k \rangle = \frac{2M}{N}$). The evolution implies new cables and costs. For the Random Graph, small-world, preferential attachment and R-MAT models, we consider an evolution in the magnitude of average node degree of $\langle k \rangle \approx 2$ then $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$. For models that do not allow explicit setting of both *size* and *order*, we operate on other parameters available that generate comparable networks. Another essential assumption is the interpretation of all the graphs generated as undirected graphs. Some of the models produce directed graph (e.g., Copying Model) since they tend to model the WWW network because of the unidirectionality of the links in the Web pages. The power grid graphs are, on the other hand, considered as undirected graphs (e.g., [4, 222, 218, 159]) since the power can flow in both directions (in principle) in the power lines that represent the edges, this aspect is even more emphasized by the vision of a prosumer-based energy exchange. A more thorough discussion concerning the other parameters required by the synthetic network models is provided in Appendix B.

Network type	Model	*Order*	*Size*	Avg. deg.	CPL	CC	Removal robustness (Rob_N)	Redundancy cost ($APL_{10^{th}}$)
LV-Small	Small-world	20	20	2.000	4.053	0.00000	0.330	7.580
LV-Medium	Small-world	90	90	2.000	11.820	0.01593	0.167	12.932
LV-Large	Small-world	200	201	2.010	17.397	0.01083	0.109	21.544
MV-Small	Small-world	250	250	2.000	24.237	0.00000	0.087	24.534
MV-Medium	Small-world	500	501	2.004	28.084	0.00000	0.057	35.413
MV-Large	Small-world	1000	1001	2.002	47.077	0.00000	0.040	60.074
LV-Small	Preferential attachment	20	19	1.900	2.579	0.00000	0.349	2.800
LV-Medium	Preferential attachment	90	89	1.978	4.315	0.00000	0.263	4.471
LV-Large	Preferential attachment	200	199	1.990	6.523	0.00000	0.206	6.375
MV-Small	Preferential attachment	250	249	1.992	5.426	0.00000	0.245	5.570
MV-Medium	Preferential attachment	500	499	1.996	5.705	0.00000	0.231	5.745
MV-Large	Preferential attachment	1000	999	1.998	6.976	0.00000	0.187	6.908
LV-Small	Random Graph	17	21	2.471	2.938	0.07451	0.390	7.472
LV-Medium	Random Graph	78	92	2.359	5.987	0.03547	0.418	10.974
LV-Large	Random Graph	172	207	2.407	6.254	0.00736	0.354	10.796
MV-Small	Random Graph	224	259	2.313	7.269	0.00000	0.322	12.002
MV-Medium	Random Graph	435	516	2.372	8.380	0.00138	0.321	12.818
MV-Large	Random Graph	863	1026	2.378	9.061	0.00070	0.328	13.446
LV-Small	R-MAT	27	31	2.296	3.615	0.00000	0.356	7.830
LV-Medium	R-MAT	88	125	2.841	4.115	0.05688	0.369	6.418
LV-Large	R-MAT	199	261	2.623	5.495	0.00737	0.364	8.774
MV-Small	R-MAT	195	263	2.697	5.629	0.00865	0.378	8.642
MV-Medium	R-MAT	365	523	2.866	5.470	0.01360	0.396	7.646
MV-Large	R-MAT	728	1056	2.901	5.726	0.00589	0.363	7.887

Table 4.3: Metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 2.

Network type	**Model**	*Order*	*Size*	**Avg. betweenness**	**Avg. betw/order**	**Coeff. variation**
LV-Small	Small-world	20	20	62.300	3.115	0.804
LV-Medium	Small-world	90	90	985.956	10.955	1.307
LV-Large	Small-world	200	201	3429.720	17.149	1.260
MV-Small	Small-world	250	250	5881.296	23.525	1.598
MV-Medium	Small-world	500	501	13980.228	27.960	1.745
MV-Large	Small-world	1000	1001	47919.616	47.920	2.279
LV-Small	Preferential attachment	20	19	31.400	1.570	2.344
LV-Medium	Preferential attachment	90	89	293.400	3.260	3.068
LV-Large	Preferential attachment	200	199	1089.260	5.446	3.288
MV-Small	Preferential attachment	250	249	1096.144	4.385	3.972
MV-Medium	Preferential attachment	500	499	2401.680	4.803	5.049
MV-Large	Preferential attachment	1000	999	6061.288	6.061	6.240
LV-Small	Random Graph	17	21	31.059	1.827	1.157
LV-Medium	Random Graph	78	92	408.308	5.235	1.126
LV-Large	Random Graph	172	207	938.512	5.456	1.276
MV-Small	Random Graph	224	259	1474.143	6.581	1.265
MV-Medium	Random Graph	435	516	3415.890	7.853	1.204
MV-Large	Random Graph	863	1026	7081.119	8.205	1.264
LV-Small	R-MAT	27	31	70.593	2.615	1.320
LV-Medium	R-MAT	88	125	282.500	3.210	1.540
LV-Large	R-MAT	199	261	937.578	4.711	1.297
MV-Small	R-MAT	195	263	959.118	4.919	1.395
MV-Medium	R-MAT	365	523	1692.910	4.638	1.581
MV-Large	R-MAT	728	1056	3633.473	4.991	2.004

Table 4.4: Betweenness metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 2.

Network type	**Model**	***Order***	***Size***	**Avg. deg.**	**CPL**	**CC**	**Removal robust-ness (Rob_N)**	**Redundancy cost ($APL_{10^{th}}$)**
LV-Small	Small-world	20	39	3.900	2.289	0.26000	0.721	4.720
LV-Medium	Small-world	90	177	3.933	3.652	0.14646	0.780	6.032
LV-Large	Small-world	200	399	3.990	4.407	0.15367	0.767	6.631
MV-Small	Small-world	250	498	3.984	4.566	0.12581	0.779	6.836
MV-Medium	Small-world	500	1000	4.000	5.067	0.10681	0.764	7.231
MV-Large	Small-world	1000	1998	3.996	5.749	0.10879	0.781	7.910
LV-Small	Preferential attachment	20	37	3.700	2.263	0.47341	0.554	4.380
LV-Medium	Preferential attachment	90	177	3.933	2.910	0.11216	0.426	4.788
LV-Large	Preferential attachment	200	397	3.970	3.322	0.09566	0.448	5.047
MV-Small	Preferential attachment	250	497	3.976	3.504	0.08400	0.419	4.998
MV-Medium	Preferential attachment	500	997	3.988	3.687	0.03929	0.401	5.232
MV-Large	Preferential attachment	1000	1997	3.994	4.211	0.01536	0.401	5.678
LV-Small	Random Graph	20	40	4.000	2.079	0.17667	0.733	4.350
LV-Medium	Random Graph	87	180	4.138	3.174	0.03418	0.735	5.368
LV-Large	Random Graph	199	400	4.020	3.869	0.03064	0.734	6.107
MV-Small	Random Graph	247	500	4.049	4.057	0.01681	0.740	6.432
MV-Medium	Random Graph	494	1000	4.049	4.495	0.00823	0.749	6.670
MV-Large	Random Graph	987	2001	4.055	5.062	0.00359	0.738	7.150
LV-Small	R-MAT	30	59	3.933	2.517	0.27360	0.579	4.511
LV-Medium	R-MAT	105	250	4.762	3.019	0.13039	0.581	4.490
LV-Large	R-MAT	227	504	4.441	3.619	0.04683	0.601	5.302
MV-Small	R-MAT	230	496	4.313	3.736	0.02940	0.626	5.381
MV-Medium	R-MAT	420	1004	4.781	3.915	0.00450	0.591	5.249
MV-Large	R-MAT	932	2039	4.376	4.562	0.00875	0.690	6.251

Table 4.5: Metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 4.

Network type	**Model**	***Order***	***Size***	**Avg. betweenness**	**Avg. betw/order**	**Coeff. variation**
LV-Small	Small-world	20	39	24.900	1.245	0.654
LV-Medium	Small-world	90	177	235.244	2.614	0.653
LV-Large	Small-world	200	399	683.780	3.419	0.703
MV-Small	Small-world	250	498	897.568	3.590	0.653
MV-Medium	Small-world	500	1000	2043.600	4.087	0.706
MV-Large	Small-world	1000	1998	4762.808	4.763	0.677
LV-Small	Preferential attachment	20	37	23.100	1.155	1.505
LV-Medium	Preferential attachment	90	177	170.644	1.896	2.219
LV-Large	Preferential attachment	200	397	463.060	2.315	2.733
MV-Small	Preferential attachment	250	497	611.520	2.446	3.017
MV-Medium	Preferential attachment	500	997	1342.864	2.686	3.484
MV-Large	Preferential attachment	1000	1997	3179.750	3.180	3.450
LV-Small	Random Graph	20	40	23.600	1.180	0.807
LV-Medium	Random Graph	87	180	196.345	2.257	0.850
LV-Large	Random Graph	199	400	589.849	2.964	0.889
MV-Small	Random Graph	247	500	766.389	3.103	0.857
MV-Medium	Random Graph	494	1000	1768.757	3.580	0.972
MV-Large	Random Graph	987	2001	4068.393	4.122	0.942
LV-Small	R-MAT	30	59	44.000	1.467	1.342
LV-Medium	R-MAT	105	250	223.733	2.131	1.695
LV-Large	R-MAT	227	504	609.419	2.685	1.493
MV-Small	R-MAT	230	496	650.374	2.828	1.468
MV-Medium	R-MAT	420	1004	1285.786	3.061	1.652
MV-Large	R-MAT	932	2039	3422.348	3.672	1.506

Table 4.6: Betweenness metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 4.

4.2.2 Model generation

Each network model described is generated and analyzed according to the significant power grid metrics described in Section 4.1. We begin with the models for

Network type	Model	*Order*	*Size*	Avg. deg.	CPL	CC	Removal robustness (Rob_N)	Redundancy cost ($APL_{10^{th}}$)
LV-Small	Small-world	20	59	5.900	1.816	0.33250	0.775	3.470
LV-Medium	Small-world	90	266	5.911	2.809	0.20131	0.794	4.508
LV-Large	Small-world	200	598	5.980	3.324	0.13596	0.797	4.895
MV-Small	Small-world	250	747	5.976	3.486	0.14477	0.798	5.039
MV-Medium	Small-world	500	1494	5.976	3.968	0.14477	0.799	5.518
MV-Large	Small-world	1000	2996	5.992	4.429	0.14854	0.797	5.905
LV-Small	Preferential attachment	20	54	5.400	1.868	0.34839	0.749	3.460
LV-Medium	Preferential attachment	90	264	5.867	2.466	0.16601	0.742	3.933
LV-Large	Preferential attachment	200	594	5.940	2.854	0.08772	0.671	4.130
MV-Small	Preferential attachment	250	744	5.952	2.926	0.08676	0.705	4.257
MV-Medium	Preferential attachment	500	1495	5.980	3.185	0.05017	0.667	4.481
MV-Large	Preferential attachment	1000	2994	5.988	3.487	0.03335	0.679	4.664
LV-Small	Random Graph	20	60	6.000	1.684	0.29599	0.775	3.370
LV-Medium	Random Graph	90	270	6.000	2.640	0.06987	0.791	4.298
LV-Large	Random Graph	200	600	6.000	3.141	0.03991	0.777	4.693
MV-Small	Random Graph	249	750	6.024	3.230	0.01934	0.793	4.884
MV-Medium	Random Graph	499	1500	6.012	3.620	0.00976	0.792	5.284
MV-Large	Random Graph	998	3000	6.012	4.022	0.00544	0.791	5.662
LV-Small	R-MAT	32	87	5.438	2.194	0.21179	0.760	3.945
LV-Medium	R-MAT	123	374	6.081	2.926	0.08173	0.717	4.377
LV-Large	R-MAT	249	759	6.096	3.165	0.04444	0.736	4.622
MV-Small	R-MAT	236	747	6.331	3.143	0.04982	0.746	4.389
MV-Medium	R-MAT	466	1512	6.489	3.427	0.04365	0.743	4.805
MV-Large	R-MAT	925	3035	6.562	3.742	0.02560	0.723	4.925

Table 4.7: Metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 6.

Network type	**Model**	***Order***	***Size***	**Avg. betweenness**	**Avg. betw/order**	**Coeff. variation**
LV-Small	Small-world	20	39	15.800	0.790	0.581
LV-Medium	Small-world	90	177	163.778	1.820	0.555
LV-Large	Small-world	200	399	464.330	2.322	0.617
MV-Small	Small-world	250	498	621.488	2.486	0.609
MV-Medium	Small-world	500	1000	1479.404	2.959	0.565
MV-Large	Small-world	1000	1998	3441.742	3.442	0.564
LV-Small	Preferential attachment	20	37	15.900	0.795	1.292
LV-Medium	Preferential attachment	90	177	133.378	1.482	2.640
LV-Large	Preferential attachment	200	397	374.970	1.875	2.401
MV-Small	Preferential attachment	250	497	485.352	1.941	2.514
MV-Medium	Preferential attachment	500	997	1095.116	2.190	2.894
MV-Large	Preferential attachment	1000	1997	2447.594	2.448	3.283
LV-Small	Random Graph	20	40	14.700	0.735	0.662
LV-Medium	Random Graph	87	180	151.489	1.683	0.809
LV-Large	Random Graph	199	400	431.090	2.155	0.835
MV-Small	Random Graph	247	500	563.839	2.264	0.710
MV-Medium	Random Graph	494	1000	1328.405	2.662	0.745
MV-Large	Random Graph	987	2001	3051.922	3.058	0.771
LV-Small	R-MAT	30	59	38.000	1.188	0.989
LV-Medium	R-MAT	105	250	247.008	2.008	1.351
LV-Large	R-MAT	227	504	550.538	2.211	1.352
MV-Small	R-MAT	230	496	530.093	2.246	1.357
MV-Medium	R-MAT	420	1004	1169.382	2.509	1.506
MV-Large	R-MAT	932	2039	2599.496	2.810	1.731

Table 4.8: Betweenness-related metrics for small-world, preferential attachment, Random Graph and R-MAT models with average node degree ≈ 6.

which it is possible to explicitly assign *order* and *size* (or one of these quantities and the average node degree); we then proceed analyzing the other models that do not explicitly allow to set the average node degree parameter.

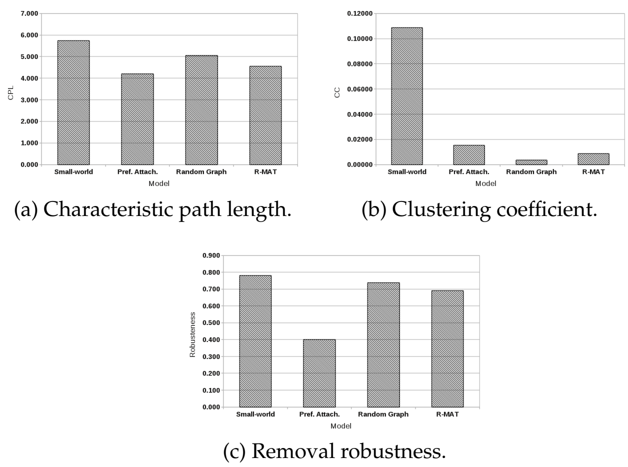
Comparison of models with average node degree $\langle k \rangle \approx 2$

The results for the metrics with average degree $\langle k \rangle \approx 2$ for the small-world, preferential attachment, Random Graph and R-MAT models score quite poorly, cf. Table 4.3. Low values for the metrics are due to the small connectivity the networks show. Especially, we highlight the low results of the small-world model under these conditions.

The betweenness analysis, whose results are presented in Table 4.4, shows an average for each node that increases with the size of the graph.

Comparison of models with average node degree $\langle k \rangle \approx 4$

Table 4.5 shows the results for small-world, preferential attachment, Random Graph and R-MAT models with an average degree $\langle k \rangle \approx 4$. One notices that the scores for the metrics improve compared to the $\langle k \rangle \approx 2$ case. The average over the characteristic path length of all the samples reduces from around 10 to a value that is slightly less than 5. The clustering coefficient has values that are significant and all positive. The small-world model scores best in this specific metric since it relies on the lattice topology that with an average degree of 4 connects each node with 4 neighbors. In particular 3 triangle structures emerge in each neighborhood of a node (of course before the rewiring process takes place). This provides a substantial contribution to the quite high clustering coefficient. A graphical comparison for the *large* sample for medium voltage considering characteristic path length, clustering coefficient and robustness are shown in Figure 4.3.

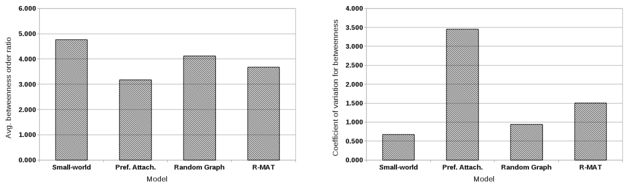


(a) Characteristic path length. (b) Clustering coefficient.

(c) Removal robustness.

Figure 4.3: Results for metrics for the *large* sample of medium voltage network type with average node degree ≈ 4.

Analyzing the betweenness we see a general improvement in the metrics compared to the $\langle k \rangle \approx 2$ case, cf. Table 4.6. The most important improvement is for the small-world model which, with approximately 4 connections per node, substantially reduces the average betweenness by a factor of 10 compared to the $\langle k \rangle \approx 2$ case. A graphical comparison for the results of the *large* sample for medium voltage type considering average betweenness *order* ratio and coefficient of variation metrics are shown in Figure 4.4.



(a) Betweenness to *order* ratio. (b) Betweenness coefficient of variation.

Figure 4.4: Results for metrics for the *large* sample of medium voltage network type with average node degree ≈ 4.

Comparison of models with average node degree $\langle k \rangle \approx 6$

Table 4.7 shows the results for small-world, preferential attachment, Random Graph and R-MAT models with an average degree $\langle k \rangle \approx 6$. The scores for the metrics considered improve even more with respect to those of Tables 4.3 and 4.5. The characteristic path length of all the samples has reduced to a value that, considering the average over all the samples with $\langle k \rangle \approx 6$, is about 3; yet 2 hops lower than the situation with $\langle k \rangle \approx 4$. The same tendency for clustering coefficient found for samples in Table 4.5 applies to this situation too. The small-world model scores highest since the neighbors of a node have 9 connections with each other (before rewiring), thus contributing to a high coefficient value.

Having increased the average degree to 6 brings benefits to the betweenness statistics too, cf. Table 4.8. The improvement on the average betweenness *order* ratio are about 25% higher than the $\langle k \rangle \approx 4$ situation; this ratio, therefore, is now very close to the experimental values that have been found for the Internet (i.e., ≈ 2.5).

A more thorough analysis of these results and a detailed comparison are available in [160].

Models Independent from the Average Node Degree

The Copying, Forest Fire, and Kronecker models are not generated using explicitly the average node degree, cf. Appendix B. Therefore, we consider the power grid metrics on them separately. We remark however that, though not explicitly used as input parameter, the average node degree of the generated graphs has similar values to those of Random Graphs, small-world and preferential attachment models generated with the same order. As a general consideration for these models, we see a reduced performance in the satisfaction of the metrics. The only models that have overall scores that become closer to the target parameters requirements is the Forest Fire and the power-law-based graph with small value of the characteristic parameter (γ). Therefore, we give a detailed analysis of these models only, and refer to [160] for the analysis of the other models.

For the Forest Fire model, we assign different forward and backward burning probabilities to obtain values for the average degree to some extent comparable with the other models. The model with $p_{fwd} = p_{bwd} = 0.2$ can be compared to models with $\langle k \rangle \approx 2$. The Forest Fire scores definitely better than all the others in clustering coefficient. This is not surprising, if one recalls the algorithm behind the model: an ambassador node is chosen and with a certain probability a certain number of ambassador's neighbors nodes are chosen to establish link to. One can see how many triangle-like structures tend to appear from such a generating method. The same observations can be done for the Forest Fire with $p_{fwd} = p_{bwd} = 0.3$ when compared to models with $\langle k \rangle \approx 4$: the characteristic path length has a score similar to node degree dependent models, while this model suffers deeply in the robustness metric which for the biggest samples obtain a score which is half compared to the other generating models with $\langle k \rangle \approx 4$. This is due to the very high damages imposed to network connectivity when high degree nodes are removed: for the biggest sample (order of about 1000 nodes), when the 20% of nodes with highest degree are removed, the biggest connected component is just 2% of the original graph *order*. This is typical of heavy-tailed distributions that Forest Fire empirically models [118]. The metric that scores best is again the clustering coefficient that is three times higher (for the biggest sample) than the already quite high value of the small-world model. Even when we consider denser Forest Fire networks (i.e., $p_{fwd} = p_{bwd} = 0.35$) the comparison with the model with $\langle k \rangle \approx 6$ brings to the same conclusions: far better clustering coefficient, but an important weakness to node removal. Betweenness for the Forest Fire model shows a trend when varying the average node degree, the more the networks becomes connected the better the metrics related to betweenness become. For the samples with a burning probability of $p_{fwd} = p_{bwd} = 0.35$, the betweenness to *order* ratio stays below 3. The same behavior applies to the coefficient of variation, although it generally scores worse than the samples already analyzed with similar average degree.

Considering the results of Random Graph with power-law models, there is a difference for the networks generated with smaller γ parameters (i.e., medium and low voltage Dutch grid $\gamma \approx 2$ and social and technological networks $\gamma \approx 2.3$) which score better than the ones with higher γ (i.e., U.S. Eastern Interconnect and Western Grid $\gamma \approx 3$ and U.S. Western power grid $\gamma \approx 4$). Refer to Appendix B for the motivations of choosing the specific γ parameters. The samples with $\gamma \approx 2$ and $\gamma \approx 2.3$ show a denser network with higher average node degree, almost double compared to the other set ($\gamma \approx 3$ and $\gamma \approx 4$); it results in a beneficial behavior for the metrics computed which present a smaller characteristic path length. This set of networks with small γ is comparable for the characteristic path length property to the values obtained for networks generated with $\langle k \rangle \approx 4$. The second set of samples (i.e., higher γ parameter) shows results that are similar to the ones obtained for samples generated with $\langle k \rangle \approx 2$. A general property that applies to all these power-law based samples is the problem they suffer from targeted attacks involving the nodes with high degree, which justifies very poor scores for robustness metric. The betweenness analysis for the power-law based models shows an average betweenness value that is smaller for the networks with a lower value for the γ coefficient so that they score very good in the betweenness to *order* ratio. A lower γ implies a higher probability in the presence of nodes that have higher node degree; usually there is quite a good positive correlation between the node degree and the betweenness the nodes have to sustain (high degree implies high betweenness for that node). It is therefore understandable why the coefficient of variation is higher for the networks characterized by a low γ than the ones with higher power-law characteristic parameter.

Comparing the generated topologies with the physical ones

The analysis of the Northern Netherlands grid shows an average degree almost constant of about $\langle k \rangle \approx 2$. In terms of average node degree the situation is similar to the high voltage grid based on the data of Eastern and Western high voltage U.S. power grid. Therefore, it is fair to compare the generated models with similar average degree, the Copying Model ones and the Random Graphs with power-law in node degree distribution with average node degree around $\langle k \rangle \approx 2$. Generated models, except the model based on Random Graph with power-law, score better than the physical topologies for all the metrics considered; the characteristic path length scores half for the R-MAT and Copying Model cases in comparison to the real data. In addition, the synthetic networks are more robust than the real data samples: R-

MAT and Random Graph score constantly above 0.3 for robustness metric while real data hardly obtain this value. Clustering coefficients are quite similar since in this configuration with limited connectivity having triangle structures in the network is rare; however, we see that R-MAT model has almost always significant clustering coefficient values. An exception is the small-world model which scores almost always worse than the real data samples. In fact, under this situation of small average node degree it is actually not fully correct to consider this synthetic topology a "small-world". The same sort of considerations apply to betweenness values: except the small-world model all the other synthetic ones score better for the average betweenness to *order* ratio metric, while for the coefficient of variation the situation is similar. If one considers the satisfaction of the desiderata for the actual samples of the Dutch medium and low voltage grid, summarized in Table 4.9, we notice that none of the parameters are satisfied. However, networks generated according to the models with almost the same average node degree (networks with $\langle k \rangle \approx 2$ in Tables 4.10 and networks based on Random Graph with power-law based on data from Eastern and Western high voltage U.S. power grid and the U.S. Western high voltage power grid in Table 4.14) do not satisfy all the desiderata as well. Therefore, this highlights that the first ingredient for the next generation of grids suited to enable local energy exchange according to the metrics defined is an increase connectivity.

Desiderata	**Northern Netherlands medium and low voltage samples**
Modularity	✗
$CPL \leqslant ln(N)$	✗
$CC \geqslant 5 \times CC_{RG}$	✗
$\overline{\upsilon} = \frac{\sigma}{N} \approx 2.5$	✗
$c_v \leqslant 1$	✗
$Rob_N \geqslant 0.45$	✗
$APL_{10^{th}} \leqslant 2 \times CPL$	$\approx$

Table 4.9: Desiderata parameter compliance of physical samples of the Northern Netherlands grid.

Increasing the average node degree naturally provides for better values for the network metrics, as shown in Tables 4.10 and 4.11. The case of the small-world model is emblematic. The $\langle k \rangle \approx 2$ case scores extremely poorly as there are not enough "shortcuts" in the network so that they can not improve much the characteristic path length. Actually, under such small average degree the condition Watts and Strogatz impose for their small-world model is not completely satisfied (i.e., $n \gg k \gg ln(n) \gg 1$, where k is the average node degree and n is the *order* of the graph). When we move closer to satisfying the small-world condition by increasing the average node degree, the value of the metrics suddenly change and the models score extremely high. The small-world scores best for the clustering property and resilience to failures in the $\langle k \rangle \approx 4$ situations. Under these conditions also betweenness values are quite concentrated around the mean with a coefficient of variation not exceeding the unit.

Comparing the average values of some metrics for the generated models while increasing node degree, one notices a natural improvement of the metrics, cf. Table 4.12. In fact, we have a reduction in characteristic path length of about 60% and an increase in the clustering coefficient of one order of magnitude; at the same time the robustness doubles. With $\langle k \rangle \approx 6$ the improvement in the metrics is less prominent, being between 10% and 20% compared to the $\langle k \rangle \approx 4$ case. From the comparison of the metric results in Tables 4.10 and 4.11, one sees that the small-world model almost always satisfies the desired requirement from a quantitative point of view when the average node degree is at least 4. From a qualitative point of view, the small-world model shows some modularity being generated starting from a regular lattice and then rewiring a certain fraction of the edges.

The models independent form average node degree perform generally worse than the other models. The adherence to the target values are shown in Tables 4.13 and 4.14. There is a general prevalence of requirement dissatisfaction, especially for parameters involving betweenness.

From the topological analysis one can see that between the models analyzed when there is a minimal connectivity ($\langle k \rangle \approx 4$ or $\langle k \rangle \approx 6$) the small-world stands out, cf. Tables 4.10 and 4.11. In Table 4.15 the models with explicit dependence on node degree are once again compared by assigning a "tick" sign (✓) for the fulfillment of each of the following properties: qualitative topological parameters (i.e., modularity), quantitative topological parameters (Tables 4.10 and 4.11) and the thrift in network realization (e.g., addition of cables which represent a cost). The latter is just an estimation, a more detailed analysis of cost in realizing a network belonging to medium and low voltage with a certain size (i.e., *small*, *medium* or *large*) and the economic benefits in electricity distribution arising from the enhanced connectivity is provided in Section 4.2.3. From Table 4.15, we conclude that networks generated with the small-world model with average degree $\langle k \rangle \approx 4$ provide the best balance between modularity, performance and, thrift for the future power grid.

4.2.3 Economic considerations

Traditionally, the problem of evaluating the expansion of an electrical system is a complex task that involves both the use of modeling, usually based on optimization techniques and linear programming [74], and the experience and vision of experts

Desiderata	**Average node degree $\langle k \rangle \approx 2$**				**Average node degree $\langle k \rangle \approx 4$**			
	SW	Pref. Attach.	Rnd. Graph	R-MAT	SW	Pref. Attach.	Rnd. Graph	R-MAT
Modularity	≈	✗	✗	✓	≈	✗	✗	✓
$CPL \leqslant ln(N)$	✗	≈	≈	✓	✓	✓	✓	✓
$CC \geqslant 5 \times CC_{RG}$	✗	✗	N/A	≈	✓	✓	N/A	✗
$\overline{v} = \frac{\sigma}{N} \approx 2.5$	✗	✗	✗	✗	✗	≈	✗	✗
$c_v \leqslant 1$	✗	✗	✗	✗	✓	✗	✓	✗
$Rob_N \geqslant 0.45$	✗	✗	✗	✗	✓	✓	✓	✓
$APL_{10^{th}} \leqslant 2 \times CPL$	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.10: Desiderata parameter compliance of the generated models with node degree $\langle k \rangle \approx 2, 4$.

Desiderata	**Average node degree** $\langle k \rangle \approx 6$			
	SW	Pref. Attach.	Rnd. Graph	R-MAT
Modularity	$\approx$	✗	✗	✓
$CPL \leqslant ln(N)$	✓	✓	✓	✓
$CC \geqslant 5 \times CC_{RG}$	✓	✓	N/A	$\approx$
$\overline{v} = \frac{\sigma}{N} \approx 2.5$	$\approx$	✓	$\approx$	✓
$c_v \leqslant 1$	✓	✗	✓	✗
$Rob_N \geqslant 0.45$	✓	✓	✓	✓
$APL_{10^{th}} \leqslant 2 \times CPL$	✓	✓	✓	✓

Table 4.11: Desiderata parameter compliance of the generated models with node degree $\langle k \rangle \approx 6$.

Avg. node degree transition	Average metric improvement (%)		
	CPL	CC	Robustness
$\langle k \rangle \approx 2 \rightarrow \langle k \rangle \approx 4$	61.7	941.6	128.5
$\langle k \rangle \approx 4 \rightarrow \langle k \rangle \approx 6$	18.0	11.8	19.6

Table 4.12: Comparison of generated topologies for varying average node degree.

in the field supported by computer computation in their decisions. With more distributed generating facilities of small capacity (e.g., rooftop photovoltaic panels), traditional methods have limits and need to be modified or updated to take into account the new scenario the smart grid framework brings into play. The models that we have so far analyzed as being candidates for the vision of the future smart grid need also to be evaluated from the economic point of view. How much will it cost to generate electrical infrastructures according to these models? What is the actual cost of adding a physical edge to the topology?

The cost of adding edges

One important difference that a physical infrastructure such as the power grid has, compared to the WWW or social networks, is the physical presence of cables that have to connect the medium voltage substations or low voltage end users generating units. If establishing a link from a Web page to another one is free, each increase in connectivity in the power grid implies costs in order to build or adapt the substation or end user premise involved and the cables required for the connection. To assess these costs in the medium and low voltage infrastructure, we consider a simple relation where the cost of cabling and cost of substations are added:

$$C_{imp} = \sum_{j=1}^{N} Ssc_j + \sum_{i=1}^{M} Cc_i \tag{4.1}$$

where C_{impl} stands for cost for implementation, Ssc_j is the adaptation cost for the substation j and Cc_i is the cost for the cable i. The cost of the cable can be expressed as a linear function of the distance the cable i covers: $Cc_i = C_{uc} \cdot l$ where C_{uc} is the cable cost per unit of length and l is the lengths of the cable. Several types of cables exist which are used for power transmission and distribution with varying physical characteristics and costs; in addition, also the cost for installation can vary significantly [148]. In the present work, to provide an initial estimate, we simply consider cabling costs and ignore substation ones. While the former are directly tied to the topology and length of the links, the latter is dependent on other factors too (e.g., type of equipment). As a source of data for cable type and pricing, we have

Desiderata	Copying Model	Forest Fire (pb=0.2)	Forest Fire (pb=0.3)	Forest Fire (pb=0.35)	Kronecker (PG params)	Kronecker (social net params)
Modularity	≈	✗	✗	✗	✓	✓
$CPL \leqslant ln(N)$	✓	✗	≈	✓	✗	✓
$CC \geqslant 5 \times CC_{RG}$	✗	✓	✓	✓	✗	✗
$\overline{v} = \frac{\sigma}{N} \approx 2.5$	≈	✗	✗	≈	✗	✗
$c_v \leqslant 1$	✗	✗	✗	✗	✗	✗
$Rob_N \geqslant 0.45$	✗	✗	✗	≈	✗	✗
$APL_{10^{th}} \leqslant 2 \times CPL$	N/A	✗	✓	✓	✓	✓

Table 4.13: Desiderata parameter compliance of the generated models.

Desiderata Model	RG with power-law (social net params)	RG with power-law (East-West US HV PG params)	RG with power-law (Western US HV PG params)	RG with power-law (NL MLV PG params)
Modularity	✗	✗	✗	✗
$CPL \leqslant ln(N)$	✓	✗	✗	✓
$CC \geqslant 5 \times CC_{RG}$	✓	$\approx$	✗	✓
$\overline{v} = \frac{\sigma}{N} \approx 2.5$	✗	✗	✗	✗
$c_v \leqslant 1$	✗	✗	✗	✗
$Rob_N \geqslant 0.45$	✗	✗	✗	✗
$APL_{10^{th}} \leqslant 2 \times CPL$	✓	✓	✓	✓

Table 4.14: Desiderata parameter compliance of the generated models.

Network Model	Avg. node deg. $\langle k \rangle \approx 2$	Avg. node deg. $\langle k \rangle \approx 4$	Avg. node deg. $\langle k \rangle \approx 6$
Small-world	✓✓	✓✓✓	✓✓
Preferential Attachment	✓	✓✓	✓
Random Graph	✓	✓✓	✓
R-MAT	✓✓	✓✓	✓✓

Table 4.15: Satisfaction of modularity, performance and cabling cost for generated models.

been provided (courtesy of Enexis B.V. the Netherlands) with cables characteristics and prices for the 11 network samples belonging to the low voltage network and 12 samples belonging to the medium voltage of the Northern Netherlands, whose topological properties we have analyzed in Chapter 3.

Statistical consideration over cables' price

The length of the cables plays an important role for both total resistance and price. If one considers the correlation between the price and resistance, high values are found using Spearman's rank correlation coefficient [102], shown in Table 24 in [160]. Especially, for generating synthetic networks it is important to obtain values for both the properties of the cables that are similar to the ones used in practice. Plotting the two variables characterizing each cable one notices that the majority of the samples concentrates in the lower tails of the joint distribution. Figure 4.5 shows the relation between the price and resistance where the values concentrate in the lower corner of $price \times resistance$ plane.

In the chart in Figure 4.5, two distinct lines deviate from the low-left corner. They represent the two main types of cables used in that sample of the low voltage network to cover different distances and result in increasing price and resistance for longer lines. The problem of extracting cable properties can be however approached in another way: *evaluate for each type of cable used in a certain sample (small, medium, and large) how the length of the cables used are distributed.* In fact, given a certain type of cable and its length all other interesting properties required for our weighted topological analysis (cf. Chapter 3) are then available (i.e., cable total resistance, cable total cost and cable supported current).

A general tendency appears when fitting the distribution of lengths to cable types belonging to low voltage and medium voltage: a fast decay in lengths' probability distribution with the majority of lengths for the low voltage cables types in the order of tens of meters, and medium voltage cables about hundreds of meters. Fitting the length to a statistical probability distribution gives a good approximation for the low voltage cable lengths as exponential distributions ($y = f_X(x;\mu) = \frac{1}{\mu}e^{\frac{-x}{\mu}}$), while for medium voltage cable lengths the distribution that generally fits best the data is the generalized extreme value distribution ($y = f_X(x;k,\mu,\sigma) = \frac{1}{\sigma}(1+k\frac{x-\mu}{\sigma})^{-1-\frac{1}{k}}\exp\{-(1+k\frac{x-\mu}{\sigma})^{-\frac{1}{k}}\}$); these hypothesis are supported in the results by the tests of Kolmogorov-Smirnov [130]. A graphical representation of the length probability cumulative distribution function for a cable type of the medium voltage network is shown in Figure 4.6.

Assume that, statistically speaking, the distribution of the lengths for each type of cable in the synthetic networks are the same as in the real samples. Therefore, knowing the probability of using a certain type of cable i ($p_{cable_i} = \frac{\#cable_i}{\sum_k \#cable_k}$ where $\#cable_i$ is the number of occurrences of cable type i in a certain network sample) that has a certain cost and resistance per meter and a specific current supported, it is then possible to estimate the cables that are used in the synthetic samples together with their properties. The joint probability of the event of having a cable of length L of cable type T can be obtained by the conditional probability relationship:

$$P(L \cap T) = P(L|T)P(T) \tag{4.2}$$

where $P(L|T)$ is the probability of the event of having a cable of length L conditioned to the event of being of cable type T, and $P(T)$ is the probability of the event of having a cable of type T.

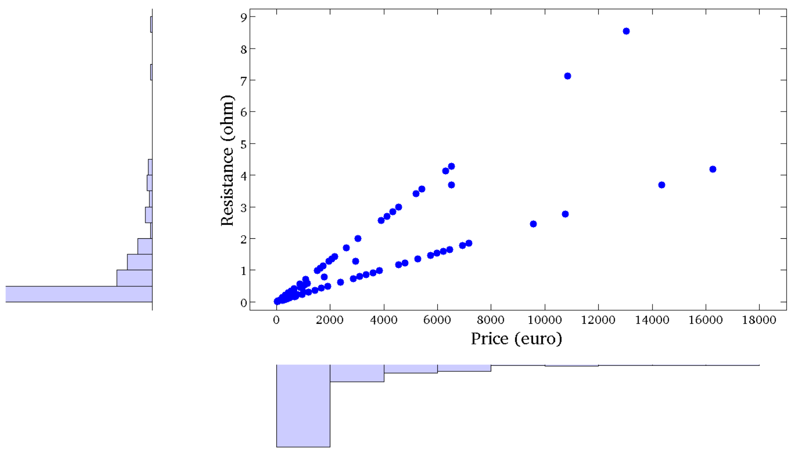


Figure 4.5: Price-Resistance pairs joint plot for the low voltage *large* size sample (i.e., LV Sample#5).

Sample type	***Size***	**Cost (thousand euro)**
low voltage - Small	≈ 20	≈ 30
low voltage - Medium	≈ 90	≈ 78
low voltage - Large	≈ 200	≈ 449
medium voltage - Small	≈ 250	≈ 32000
medium voltage - Medium	≈ 500	≈ 42000
medium voltage - Large	≈ 1000	≈ 43000

Table 4.16: Cabling cost for $\langle k \rangle \approx 2$ synthetic samples.

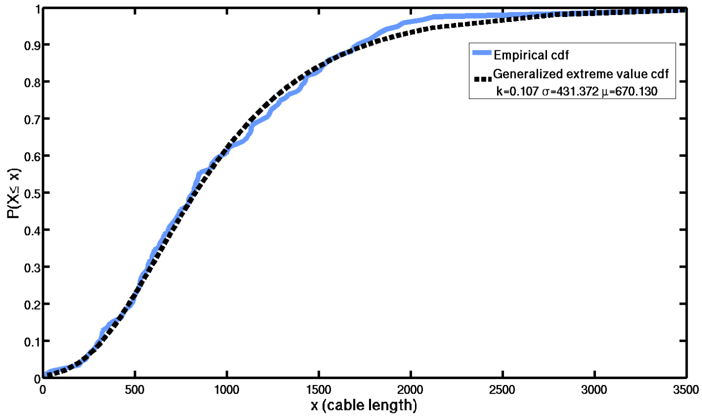


Figure 4.6: Cumulative distribution function for cable length (meters) for cable type "3x1x70al" in Northern Netherlands medium voltage sample size *medium*.

Economic benefits of highly connected topologies

Once the information about cable prices is available, it is possible to estimate the cost for realizing a network with a certain connectivity and whether such networks are able to lower the (economic) barrier towards decentralized energy trading in terms of cost of electricity distribution. The results for low voltage and medium voltage networks for *small*, *medium* and *large* types with an average node degree $\langle k \rangle \approx 2$, $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$ are shown in Table 4.16, Table 4.17, and Table 4.18 respectively.

For medium voltage networks, it is important to clarify that the information available for cables' prices in this study are only partial and limited to some technologies (only few cross sections of aluminum and copper cables). Anyway, in order to have a glimpse of costs for this type of the network, we fitted to the best interpolating curve the available prices as a function of the cross section. The relation between price and cross section for aluminum cables fits best to a cubic polyno-

Sample type	*Size*	Cost (thousand euro)
low voltage - Small	≈ 40	≈ 51
low voltage - Medium	≈ 180	≈ 174
low voltage - Large	≈ 400	≈ 827
medium voltage - Small	≈ 500	≈ 55000
medium voltage - Medium	≈ 1000	≈ 88000
medium voltage - Large	≈ 2000	≈ 86000

Table 4.17: Cabling cost for $\langle k \rangle \approx 4$ synthetic samples.

Sample type	*Size*	Cost (thousand euro)
low voltage - Small	≈ 60	≈ 76
low voltage - Medium	≈ 270	≈ 254
low voltage - Large	≈ 600	≈ 1239
medium voltage - Small	≈ 750	≈ 80000
medium voltage - Medium	≈ 1500	≈ 132000
medium voltage - Large	≈ 3000	≈ 131000

Table 4.18: Cabling cost for $\langle k \rangle \approx 6$ synthetic samples.

mial, while for the copper ones is linear; in this way we estimate the prices for all the types of cables involved knowing their cross section. The small difference in costs between the *medium* and *large* types of networks for medium voltage is related mainly to the different technologies (i.e., cable types) in the cables that are used for these types of networks according to the sample data provided.

Price alone is not enough to describe future scenarios. It is important to investigate how an enhanced connectivity is beneficial to the electricity distribution costs. We have shown the benefits in topology-related metrics for more connected networks earlier in this section, however those results consider only the topology without any parameter related to the properties of the cables (e.g., resistance and supported current). In order to consider the effects of topology in electricity distribution costs, we resort to the α and β metrics that we have proposed in Chapter 3. In order to apply these metrics to power grid networks *weights* are essential, representing physical quantities such as resistance of the cable and maximal operating current supported by the cable. Once we have the statistical information about the types and the length of the cables used in a specific type of physical network, (i.e., medium or low voltage and its *small*, *medium* or *large* size) it is possible to assign *weights* to the edges of the generated graphs. This is done under the assumption that the same type of cables are used and that the distances covered in general (i.e., statistically) remain the same.

We consider α and β for networks generated following the small-world model since it has proven to be the best one in the pure topological analysis. For low voltage networks we compute the metrics for networks with an increasing average node degree ($\langle k \rangle \approx 2$, $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$). It is not surprising to see that the samples with $\langle k \rangle \approx 2$ score poorer than the other networks. The network with *medium* size scores best and the difference between the network with $\langle k \rangle \approx 6$ and the network with $\langle k \rangle \approx 4$ is limited. Robustness (i.e., β parameter) for the *medium* and *large* size networks reaches an high value just with a sufficient connectivity (i.e., $\langle k \rangle \approx 4$) and more connectivity (i.e., $\langle k \rangle \approx 6$) does not improve much this metric. The samples with *small* size score better in the α metric and this is quite reasonable since the paths are limited, of course due to the reduced *order* of the network.

Considering α and β for the networks generated for the medium voltage, the same tendency appears: once the network is sufficiently connected (i.e., $\langle k \rangle \approx 4$) the metrics score definitely better than the $\langle k \rangle \approx 2$ situation.

Let us compare α and β of the synthetic networks with the values of the current power grid samples of the Northern Netherlands. Considering the low voltage samples and the synthetic networks designed for this purpose, we generally see an improvement in the metrics especially in the α values for the $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$ networks. In fact, if we do not consider the synthetic networks with $\langle k \rangle \approx 2$, because of the problems of small-world topology with such small connectivity, there is an improvement on average in the α metric for synthetic samples with $\langle k \rangle \approx 4$ of more than 50% compared to the Northern Netherlands samples. In fact, for the α metric from an average of about 12 for the physical samples, the $\langle k \rangle \approx 4$ synthetic ones score about 6. The improvement is more than 60% when considering the $\langle k \rangle \approx 6$ ones where the average for these synthetic networks scores just below 5. There are improvements also in the β metric, although limited. From an average around 4 for the physical samples, the $\langle k \rangle \approx 4$ ones score on average just below 2.75; while a better result is obtained by $\langle k \rangle \approx 6$ which on average score 2.30 (about 40% improvement). The graphical comparison between Dutch samples (white diamonds) and generated samples (white circles) is shown in Figure 4.7 in which each symbol represents a sample in the α, β quadratic function envelope that is chosen as the type of dependence between the topological parameters and electricity transport prices.

Taking into account the medium voltage, the Dutch samples of the Northern Netherlands and the small-world synthetic networks, we see an important improvement in the metrics both in the α and β values for the $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$ networks. As already mentioned, synthetic networks with $\langle k \rangle \approx 2$ should not be considered. The improvement on average in the α metric is more than 65% comparing to the $\langle k \rangle \approx 4$ synthetic samples (from an average of α about 33 for the physical samples, the $\langle k \rangle \approx 4$ synthetic ones score about 11), and an improvement of more than 75%

when comparing to the $\langle k \rangle \approx 6$ ones (the average α for $\langle k \rangle \approx 6$ synthetic networks scores around 7.3). There are improvements also in the β metric. In particular, from an average of around 3.55 for the physical samples, the $\langle k \rangle \approx 4$ score on average just below 1.15; a similar result is obtained by $\langle k \rangle \approx 6$ which on average score about 1.2 (more than 65% improvement). The graphical comparison is shown in Figure 4.8.

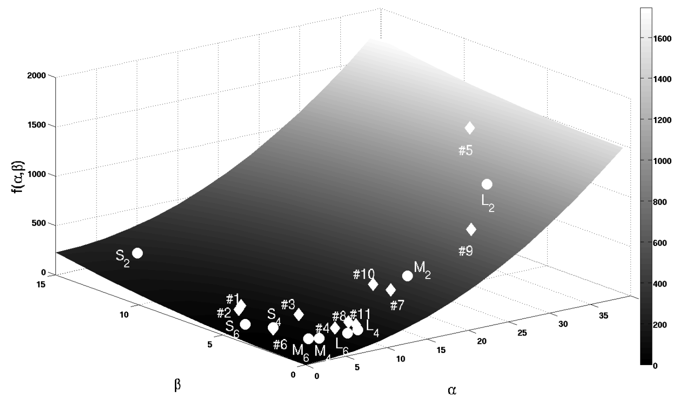


Figure 4.7: Comparison of the transport cost between synthetic small-world networks (withe circles) and Northern Netherlands low voltage samples (white diamonds).

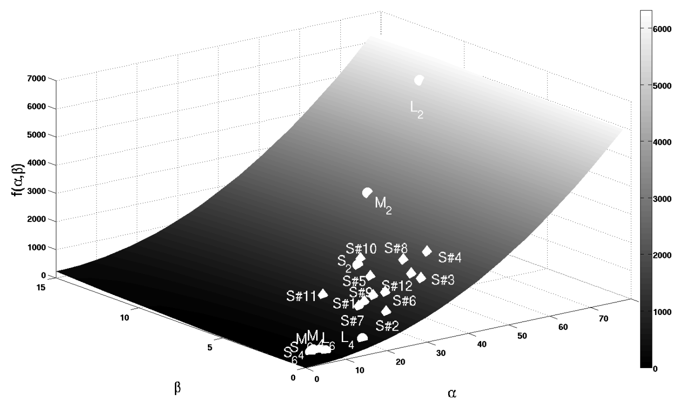


Figure 4.8: Comparison for transport cost between synthetic small-world networks (withe circles) and Northern Netherlands medium voltage samples (white diamonds).

4.2.4 Topology costs vs. benefits

Watts and Strogatz's small-world model captures best the requirements for the new grid compared to the others analyzed being these dependent on the average node degree (preferential attachment, R-MAT and Random Graph) or not (Copying Model, Forest Fire, Kronecker and power-laws). The higher clustering that this models exhibits provides efficient local distribution with paths that are locally short; at the same time the shortcuts between the local clusters are the elements that keep the (global) average path limited. These two aspects influence the α parameter which then stays relatively small. At the same time, the small-world model benefits from a general robustness against failures: the absence of big hubs that keep the network together (which are present on the other hand in the power-law-based topologies, for instance) improves the reliability against attacks, which help obtaining good scores for the β parameter. More quantitatively, one sees the general improvement in the metrics characterizing both the parameters influencing the losses (i.e., α parameter) and the reliability of the grid (i.e., β parameter) while the network becomes more dense, i.e., more edges are added. On average, we see an improvement of at least 50% when comparing the physical samples of Northern Netherlands with the small-world networks with an average degree $\langle k \rangle \approx 4$, while better results are obtained with more density (i.e., $\langle k \rangle \approx 6$) where the improvement are 60% compared to the physical samples. This is indeed beneficial to the power grid and, according to the relationship with the topology, it should translate into a reduction in the costs for electricity distribution since α and β are composed by essential ingredients influencing electricity distribution price.

These benefits come literally at a cost. The network needs more connectivity therefore costs for extra cabling need to be considered in addition to the cost for upgrading the substations and end users electricity gateways. A comprehensive return on investment analysis on this aspect is beyond the scope of the present study. Nevertheless, it is interesting to see how with the α and β metrics it is possible to consider how a certain physical sample belonging to a certain size category (*small*, *medium* and *large*) would improve in its performance if its topology is arranged according to the principles of a synthetic model and more connections are added appropriately.

The benefits reached for α and β should translate into a reduction in the cost for electricity transport and distribution since the parameters that influence these metrics are directly connected to aspects related to costs. However, the significant investment required to add more connectivity in the network might not immediately enable cheaper electricity costs, but on the contrary make it more expensive.

4.3 Evolution of Current Distribution Networks

Another important topic to consider is how to evolve or adapt the current distribution grids to support a prosumer-based energy approach. It is therefore essential to evaluate the benefits of different ways (or strategies) of evolving the distribution network in terms of additional connections. It is also essential to consider both the topological benefits together with costs of realization and reduction in cost of the electricity distribution. As a case study we consider the evolution of the distribution grid of the Netherlands that we have analyzed in Chapter 3.

4.3.1 Network evolution policies

Next, we consider evolutions starting from the Dutch samples, that is, adding cables according to several strategies of network growth. We break the evolutions into four groups of edge growth: increments of 25%, 50%, 75%, and 100%. The choice of stopping at 100% is performed based on the results of Section 4.2, where we show that an average node degree of $\langle k \rangle \approx 4$ has the right balance of improved network qualities and costs of network evolution. We consider several strategies for evolving the graph by adding more links, namely:

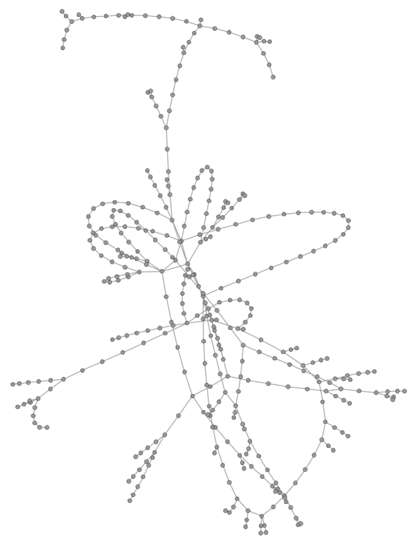
- **Assortativity.** A network is assortative if nodes having similar characteristics or properties are connected one another [150]. We consider node degree and then take two strategies:
  - *High degree* nodes are connected one another. The process starts with considering a set of nodes with the highest equal node degree and connect them together. The process goes on considering the next set of nodes with equal high degree in the order of rank and so on.
  - *Low degree* nodes are connected one another. The process goes on as for the the high degree strategy, but nodes are linked starting from the couples with lowest degree.
- **Dissortativity** is the opposite of assortativity, that is, a network is dissortative if nodes having different characteristics or properties are connected together. Therefore, nodes with highest node degree are linked to nodes with lowest node degree.
- **Triangle closure** is based on the principle of increasing the clustering coefficient of the network. At each step, a node is selected at random and for each pair of its neighbors an edge is added between them, if not already present.

- **Least distance** gives priority to the connection of nodes that are geographically closer to each other. This strategy can minimize the costs of cabling since such costs are directly proportional to the length of cables.

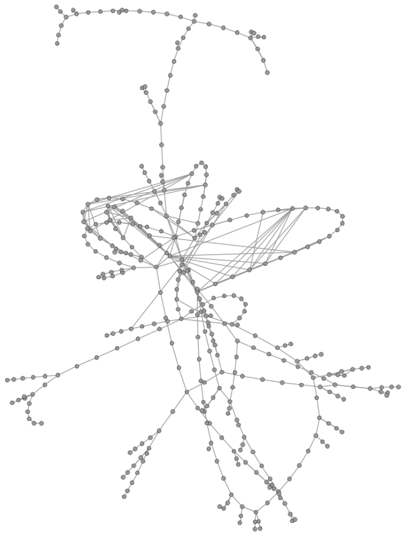
- **Random** is based on the random selection of nodes to attach edges. At each step, a pair of distinct nodes are randomly selected and an edge between them is added.

For every strategy, if two nodes already have an edge that connects them, the edge is not added and the evolution strategy continues. In fact, in the graph models, we only allow a single edge between a pair of nodes, if not already present.

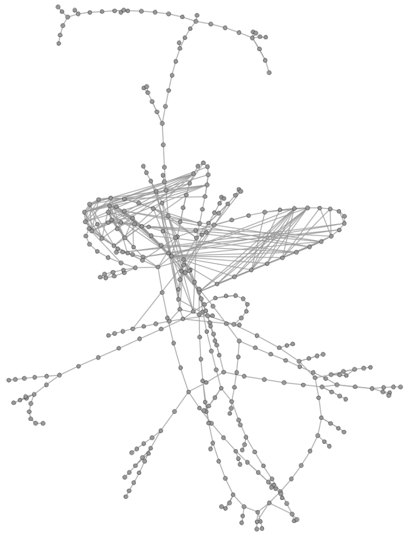
We study the adaptation of current physical networks according the strategies just described and we analyze the obtained graphs according to a set of metrics that provide a view of efficiency of the whole network and its adequacy for local energy distribution (cf. Section 4.1).



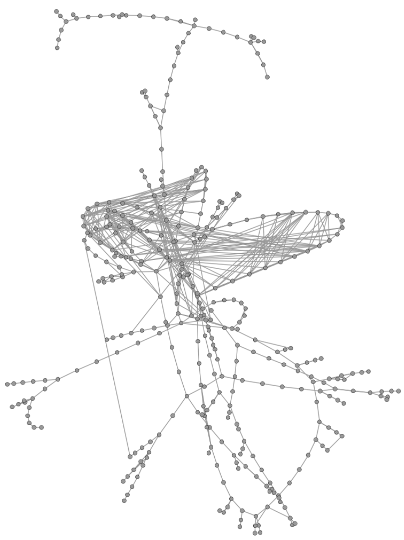
(a) Physical sample.



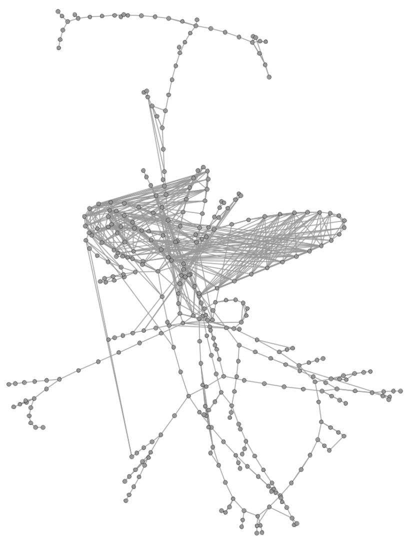
(b) 1st stage of evolution (i.e., +25% edges).



(c) 2nd stage of evolution (i.e., +50% edges).



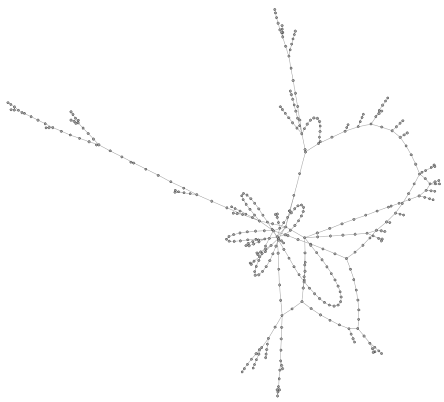
(d) 3rd stage of evolution (i.e., +75% edges).



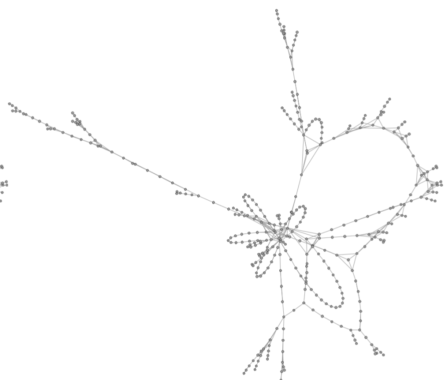
(e) 4th stage of evolution (i.e., +100% edges).

Figure 4.9: Stages of evolution of medium voltage sample #8 following least distance strategy.

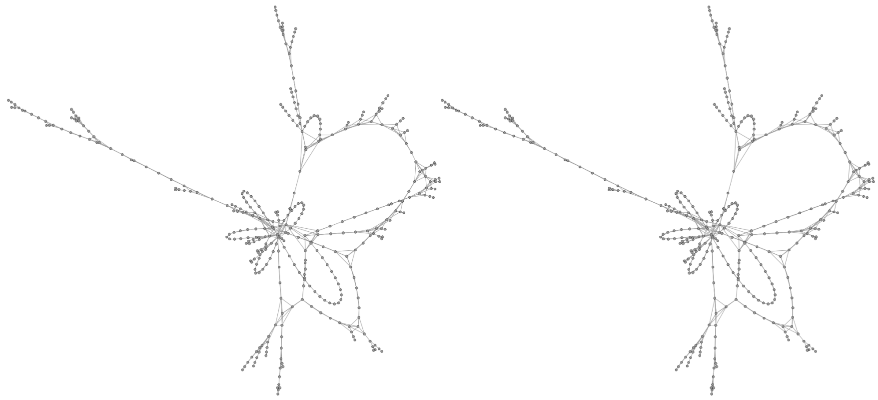
To determine which strategy is more rewarding and provides the higher benefits, one needs to compare the metrics obtained with the different evolution strategies. If determining a "best" strategy is unrealistic, one should aim at identifying trade-offs between one strategy compared to another one. We evaluate all the strategies considering average values over all the samples and considering the satisfaction of the topological metrics defined in Section 4.1. Visual examples of the evolution strategies are provided in Figures 4.9 and 4.10. Figure 4.9 represents the evolution following the least distance strategy, one can see that for a part of the network nodes that are several hops apart in the original network (they belong to two different branches) are now connected by new edges providing a substantial topological shortcut in the network. Figure 4.10 shows the evolution following the triangle closure strategy, one sees the appearance of several new triangle structures in the network.



(a) Physical sample.



(b) 1st stage of evolution (i.e., +25% edges). (c) 2nd stage of evolution (i.e., +50% edges).



(d) 3rd stage of evolution (i.e., +75% edges). (e) 4th stage of evolution (i.e., +100% edges).

Figure 4.10: Stages of evolution of medium voltage sample #8 following triangle closure strategy.

4.3.2 Comparison of the evolution strategies

Comparison of the evolution strategies for the medium voltage grid

It is difficult and perhaps even methodologically wrong to establish "a winner" among the various evolution strategies. The improvement in the various metric values during the evolution does not establish a linear pattern. As a general remark, we reinforce that the addition of links is beneficial to the topological metrics analyzed and one might speculate that the optimal solution in order to minimize (or maximize) the metrics is to realize a fully connected network (i.e., clique). With a real infrastructure such as the power grid this is impossible to achieve from the economic point of view, but also from a technical point of view (e.g., substations receiving thousand of connections).

The comparison of the averages over all samples in each evolution step for the various strategies is shown in Table 4.19. Take the triangle closure strategy, for instance, it is the one that in every step of the evolution process leads to the maximal values of the clustering coefficient, but it is also the one that has the worst results together with the least distance strategy for the metrics related to path length (characteristic path length and $APL_{10^{th}}$). Considering the robustness metric, the triangle closure offers the smallest improvements, while for the betweenness-related metrics, the triangle closure strategy is the worst compared to all the others with the second set of highest values for average betweenness and an increasing coefficient of variation. The same considerations apply to the least distance strategy that to a certain

	Evolution step	CPL	CC	Rob.	10^{th} red. path	Avg. bet.	Avg. bet./*order*	Coeff. var.
Assortative high degree	+25%	5.757	0.037	0.326	8.550	1682.845	4.744	2.258
	+50%	5.121	0.067	0.344	7.095	1369.399	4.100	2.259
	+75%	4.724	0.098	0.361	6.493	1245.963	3.799	2.147
	+100%	4.558	0.122	0.367	6.072	1226.062	3.717	1.904
Assortative low degree	+25%	6.404	0.045	0.358	9.503	1955.047	5.598	2.039
	+50%	6.320	0.090	0.408	8.492	1931.053	5.508	1.651
	+75%	6.261	0.109	0.482	8.149	1913.409	5.440	1.504
	+100%	6.172	0.121	0.496	8.014	1888.415	5.357	1.441
Triangle closure	+25%	10.761	0.165	0.257	14.559	2698.910	8.130	1.786
	+50%	9.504	0.288	0.321	11.680	2377.892	7.123	1.872
	+75%	8.482	0.388	0.341	10.053	2104.554	6.289	1.927
	+100%	7.618	0.483	0.384	8.997	1884.527	5.563	2.034
Dissortative	+25%	5.803	0.019	0.349	8.891	1727.708	4.949	2.433
	+50%	5.591	0.028	0.408	7.904	1645.824	4.710	1.953
	+75%	5.402	0.034	0.463	7.421	1594.361	4.532	1.739
	+100%	5.217	0.033	0.468	7.053	1559.650	4.418	1.613
Least distance	+25%	10.025	0.141	0.272	13.078	3044.817	9.042	1.703
	+50%	9.066	0.214	0.360	11.696	2789.684	8.127	1.744
	+75%	8.535	0.261	0.455	10.705	2646.133	7.652	1.788
	+100%	8.109	0.304	0.523	10.013	2525.651	7.275	1.807
Random	+25%	6.249	0.011	0.423	9.954	1785.946	5.265	1.194
	+50%	4.990	0.016	0.634	7.780	1378.933	3.998	1.074
	+75%	4.393	0.019	0.737	6.739	1178.359	3.394	0.964
	+100%	4.000	0.027	0.764	6.080	1044.279	2.988	0.881

Table 4.19: Comparison of evolution strategies for medium voltage network.

extent is very similar to the triangle closure. This strategy is the worst concerning the path length, having a characteristic path length more than double than the best evolution strategy (i.e., random); the same also applies for the $APL_{10^{th}}$. Concerning redundancy aspects at the end of the evolution process, the least distance strategy is the one (except for random) that has the best robustness results. In terms of metrics related to betweenness, the least distance strategy shows results that are even worse compared to the triangle closure one. Considering the high degree assortative evolution strategy the results show that, overall, it is the best evolution strategy in the first two steps of the evolution; in the first step, it is even better than the random one (with the exception of robustness). Considering betweenness metrics, this strategy is second only to the random strategy, with the exception of coefficient of variation which is worse than all other metrics. In general, this is a good strategy for evolving the network especially concerning the path-related aspects, with just robustness that lacks behind compared to other evolution strategies. The assortative low degree strategy is not particularly appealing and it is outrun by the assortative high degree strategy in every metric except robustness and coefficient of variation of betweenness, therefore we do not find it particularly interesting for evolution purposes of the medium voltage networks. The dissortative strategy does not excel particularly in any of the metrics considered, however its values are quite fair, especially in the initial stage of the evolution. In particular, since the values of the metric do not improve substantially in the following steps of the evolution process, such evolution strategy could be used as a slightly more robust alternative compared to the assortative high degree in those scenario where the number of edges to be added is minimal and there is no need of special excellence in one topological parameter. When enough connectivity is added, the evolution strategy that scores best compared to all the others for all the metrics, with the exception of the clustering coefficient metric, is the random one. It has already been noted by Casals and Murtra [184] that some randomness in the network is beneficial especially for aspects related to robustness. In our analysis, we have the same general results, with characteristic path length about 4 and a $APL_{10^{th}}$ of just 6. In addition, the networks evolved according to this strategy are the most robust with a value higher than 0.7 when the maximal connectivity is reached. The same considerations apply for betweenness metrics that obtain the best results when the connectivity is enhanced reaching a betweenness to *order* ratio smaller than 3 and a coefficient of variation below 1 at the last evolution step. However, it is difficult to propose for a distribution provider to improve his grid in a random fashion, even if part of the weaknesses and inefficiencies come from the lack of such randomness. Considering rational and evolution strategies that come with a motivation we consider the assortative high degree strategy as the one that, by evaluating the various topological metrics, scores best in the evolution tests that we have performed on the Dutch medium voltage grids.

Comparison of the evolution strategies for the low voltage distribution grid

As for the medium voltage evolution strategies, one cannot declare a winner. Also in this case some improvements between the various phases of evolution are not linear, therefore a strategy that scores best in one evolution step could result worse compared to another one when more edges are added. We point out once again that the addition of links is beneficial to the topological metrics analyzed and one might speculate that the optimal solution to minimize (or maximize) the metrics is to realize a fully connected network. With a real infrastructure such as the power grid this is impossible to achieve due to economic and technical considerations.

The comparison of the averages over all samples in each evolution step for the

	Evolution step	CPL	CC	Rob.	10^{th} red. path	Avg. bet.	Avg. bet./*order*	Coeff. var.
Assortative high degree	+25%	3.600	0.062	0.408	7.535	207.959	2.515	1.518
	+50%	3.225	0.157	0.432	5.797	191.974	2.227	1.564
	+75%	2.973	0.234	0.457	5.059	182.954	2.043	1.678
	+100%	2.729	0.277	0.481	4.977	174.593	1.830	1.772
Assortative low degree	+25%	3.866	0.057	0.405	8.865	226.416	2.869	1.436
	+50%	3.074	0.127	0.446	6.310	188.824	2.093	1.746
	+75%	2.827	0.223	0.509	5.187	177.928	1.825	2.113
	+100%	2.766	0.277	0.551	4.730	175.312	1.765	1.939
Triangle closure	+25%	5.908	0.131	0.346	9.028	488.127	4.599	1.202
	+50%	5.463	0.250	0.383	8.070	461.893	2.044	1.247
	+75%	5.020	0.355	0.456	6.973	419.114	1.850	1.310
	+100%	4.686	0.462	0.497	6.573	392.781	1.732	1.428
Dissortative	+25%	3.396	0.093	0.403	7.923	194.838	2.413	1.849
	+50%	2.826	0.177	0.435	6.168	167.339	1.875	2.318
	+75%	2.616	0.214	0.465	4.924	150.914	0.667	2.459
	+100%	2.476	0.232	0.501	4.618	140.275	0.619	2.502
Random	+25%	3.748	0.039	0.440	8.404	219.250	2.650	1.159
	+50%	3.171	0.095	0.583	6.323	172.730	2.113	1.054
	+75%	2.841	0.141	0.636	5.450	149.836	0.642	1.038
	+100%	2.621	0.156	0.689	4.831	132.417	0.567	0.978

Table 4.20: Comparison of evolution strategies for low voltage network.

various evolution strategies is shown in Table 4.20. Once again, the strategy that scores best in comparison with the others is the random one for robustness and for metrics related to betweenness. For robustness, in the last stage of evolution the value reaches on average 0.7. For betweenness-related metrics the two final stages of evolution are the best ones for average betweenness and average betweenness to *order* ratio. This strategy is the only one for which the coefficient of variation results below one (on average for all the samples) in the final stage of evolution. However, the random strategy has its weak point in the clustering coefficient metric that score below all the others evolution strategies. Considering the assortative strategies, one sees that the two strategies have quite similar scores for the metrics. In the very first stage, the high degree assortative strategy is better, but later in the evolution the assortative low degree slightly outperforms the assortative high degree one for clustering coefficient and characteristic path length, while for the robustness the difference is limited. The results between these two metrics for betweenness are quite similar and here the only interesting difference is in the coefficient of variation which scores best for the high node degree strategy. These two strategies are comparable to the random one for characteristic path length matters, while scoring worse for robustness, but better for clustering. Considering the clustering coefficient metric alone, the strategy that outperforms the others is the triangle closure strategy with values that are double compared to the others at the end of the evolution process. However, this strategy is worse than all the others for the characteristic path length. Although it has the highest values for average betweenness, the coefficient of variation ranks second compared to the other strategies considered. The strategy that slightly outperforms the others (except the values for robustness of the random one) is the dissortative strategy. In fact, the path length is just smaller than 2.5 even better than the random edge addition. The clustering coefficient is in line with the assortative strategies and scores around 0.5. Concerning betweenness, this strategy scores best considering the average betweenness, but for the coefficient of variation this strategy scores worst. In the low voltage distribution grid it is difficult to propose a strategy to follow for the evolution given the similarity for the values of the evolution strategies considered. However, considering non-random evolution strategies we consider the dissortative strategy a good candidate since scores best in the evolution tests that we have performed on the Dutch low voltage grids.

Parameters satisfaction for medium voltage networks

We now consider the satisfaction of the quantitative metrics to analyze the appropriateness of synthetic topologies to improve the grid in reducing losses, facilitate local energy distribution, and increasing network robustness. For the tables containing the numerical values of the results discussed here we refer to [164].

Assortative high degree The assortative high degree evolution strategy satisfies the desired values for the metrics concerning the characteristic path length already from the second step of evolution for ten out of twelve samples. All samples satisfy the requirements over the redundant path. Considering the clustering coefficient, the majority of samples satisfies the requirement at the third step of evolution where seven samples have a $CC \geqslant 5 \times CC_{RG}$; in the final stage of evolution this metric is satisfied by all the samples except two. The real drawback of this metric is represented by the robustness that never reaches the goal of 0.45, but stops around 0.37. The unsatisfaction of the metric is also present for betweenness ones: all samples are one unit larger than the target even in the last step of evolution. In addition, the coefficient of variation never reaches the target for the samples.

Assortative low degree The assortative low degree evolution strategy satisfies the desired values for the metrics concerning the characteristic path length already from the second step of evolution with eleven out of twelve samples. All the samples satisfy the requirements over the redundant path already from the first step. Considering the clustering coefficient, the majority of samples satisfies the requirement at the second step of evolution where seven samples have a $CC \geqslant 5 \times CC_{RG}$; in the final stage of evolution this metric is satisfied by all the samples except one. This strategy almost provides the satisfaction of the robustness requirement from the third step on by having six samples fully compliant and three very close to the 0.45 threshold. The metrics concerning betweenness are not satisfied both for betweenness to *order* ratio and for the coefficient of variation which never reach the target for the samples at any stage of evolution.

Triangle closure The triangle closure evolution strategy focuses on the clustering coefficient, therefore reaching the target already in the first step of the evolution. The evolved graphs do not only satisfy the requirement posed of having $CC \geqslant 5 \times CC_{RG}$, but also a more restrictive requirement of $CC \geqslant 10 \times CC_{RG}$ which can be considered the condition for satisfying the small-world requirement for this property cf. [223]. Concerning the path length properties, the characteristic path length is never lower than the logarithm of the *order* of the graph. The requirement over the $APL_{10^{th}}$ is satisfied already from the first addition of edges. On all other requirements this strategy is weak. For robustness just few samples reach values around 0.4, while the majority is about 0.3 when the most of the edges are added. For metrics that involve betweenness, this evolution strategy scores poorly and it is not close to the target for both average betweenness to *order* ratio and for the coefficient of variation for any single sample.

Dissortative node degree The dissortative node degree strategy scores quite poorly. As mentioned before, this strategy does not excel in one specific metric, but all the metrics are slightly improved. Such aspects result in a limited crossing of the threshold for the desiderata parameters imposed. Only after the third evolution step the characteristic path length is almost satisfied: nine of the twelve samples reach the target, while one more reaches the target when even more edges are added. A similar condition is true for the robustness metric. In fact, six samples satisfy the target after the third step and other three are above the 0.4 value. The clustering coefficient is always below the threshold, having values that often are slightly higher than the clustering coefficient of a Random Graph with same *order* and *size*. The only metric fully satisfied is the one related to redundant paths in the network. While the metrics related to betweenness are not satisfied even at later stages.

Least distance The results of the least distance strategy are similar to those of triangle closure. The clustering coefficient desiderata is reached already in the first step of the evolution. From the second step on, the evolved graphs do not only satisfy the requirement posed of having $CC \geqslant 5 \times CC_{RG}$, but also a more restrictive requirement of $CC \geqslant 10 \times CC_{RG}$ which can be considered the condition for satisfying the small-world requirement for this property cf. [223]. Concerning the path length properties, the characteristic path length is never lower than the logarithm of the *order* of the graph. However, the easy requirement related to the $APL_{10^{th}}$ is satisfied already from the first addition of edges. Robustness is another positive note of this evolution strategy. From the third step on the seven samples reach the 0.45 target and other three are close to 0.4; in the last step basically all the medium voltage evolved samples satisfy the desiderata. The situation is not so positive for betweenness related metrics since they score worse compared to the others strategies for both average betweenness to *order* ratio and for the coefficient of variation.

Random The random evolution strategy is the one that satisfies most of the desiderata parameters not only after the first step of evolution, but already after the second one. The metrics concerning the characteristic path length are satisfied basically already from the second step of evolution. The same applies to the redundant path whose goal is met already in the first evolution step. Robustness is really impressively achieved after the addition of 50% of more edges, with all the samples above the 0.45 threshold and the majority of them scoring even higher, above 0.6. The metrics related to betweenness that usually fail for the other strategies are here met completely on the coefficient of variation side after the third evolution stage (eight out the twelve samples meet the target). In addition, this evolution strategy is the one that also goes closer to the satisfaction of the betweenness to *order* ratio. The only problem of such a strategy is on the clustering coefficient side. With such a strategy the formation or closure of topological triangle structures is very unlikely.

Parameters satisfaction for low voltage networks

We now consider the satisfaction of the quantitative metrics for the low voltage evolved samples. For the tables containing the numerical values of the results discussed here we refer to [164].

Assortative high degree The assortative high degree evolution strategy almost satisfies the desired values for the metrics concerning the characteristic path length from the second step of evolution with eight out of eleven samples; while full compliance is in the last evolution step. However, the requirement for the redundant path are never fully satisfied. Considering the biggest samples, they never satisfy the condition $CC \geqslant 5 \times CC_{RG}$. Also robustness requirement are not satisfied, only six samples in the last two evolution stages reach the goal of 0.45. The betweenness requirement is only partially satisfied: eight out of the eleven samples have a betwenness to *order* ratio around the target of 2.5. Considering the coefficient of variation, only the two smallest samples satisfy this property.

Assortative low degree The assortative low degree evolution strategy satisfies or is close to satisfying the metrics in the last stage of the evolution process. For characteristic path length, already from the second step, ten of the eleven samples reach the target; on the contrary, the target is never reached for the redundant path requirement. The three biggest samples score sufficiently concerning the clustering coefficient requirement, while the others have high values, but never satisfy the high-demanding requirements. Robustness is close to satisfaction in the last stage of evolution where all samples reach the target except two that stop to a value around 0.4. The metrics concerning betweenness show an almost satisfaction of the average betweenness to *order* ratio with eight of the samples well below the target; however the coefficient of variation never reaches the target for the samples at any stage of evolution.

Triangle closure The triangle closure evolution strategy focuses on the clustering coefficient, therefore reaching the target already after the second step of the evolution. Two of the three biggest samples satisfy also the small-world requirement for this property cf. [223]. Concerning the path length properties, the characteristic path length is never lower than the logarithm of the *order* of the graph. However, the requirement over the $APL_{10^{th}}$ is satisfied already after the second step of evolution. In fact, the redundant paths are smaller than twice the characteristic path length. For robustness, seven out of the eleven samples reach the target in the last step of evolution. For metrics that involve betweenness, only seven samples are below the target for the betweenness to *order* metric, actually the smallest in *order*, while the three biggest samples are far from the target. The coefficient of variation increases in the evolution.

Dissortative node degree The dissortative node degree strategy scores quite well for evolving low voltage networks. Especially in the last stage of evolution half of the metrics are satisfied or very close to satisfaction. Already in the second step of evolution, the characteristic path length is satisfied; actually already in the first step of evolution eight samples satisfy the condition. Robustness is satisfied with nine out of the eleven samples that fully comply with the desiderata and the other two that are above 0.4. The real weak point is the clustering coefficient which even for the biggest samples is never close to the target. Betweenness-related metrics have a dissimilar behavior: the betweenness to *order* ratio satisfies the desiderata for nine samples, while the coefficient of variation has an increasing trend and in the final step does not comply with the requirements for any sample.

Random The random evolution strategy proves to be one of the strategies that satisfies most of the parameters. The metrics concerning the characteristic path length are satisfied basically already from the second step of evolution with ten of the eleven samples compliant. A different behavior applies to the redundant paths whose goal is not met, but in the final step eight of the eleven samples reach the target. Robustness is almost achieved already after the second step: all the samples score above 0.4 and just three do not reach the target in that step, but later in the evolution. In the last step, the satisfaction of the redundant path is almost satisfied with eight of the samples (containing also the three biggest ones). The metrics related to betweenness that usually fail for the other types of evolution are here almost entirely met concerning the coefficient of variation and the betweenness to *order* ratio in the last stage of evolution. The main drawback is on the clustering coefficient side. With such a strategy the formation or closure of topological triangle structures is difficult.

Discussion

Considering the medium voltage grid, the random evolution strategy is the one that satisfies most of the desiderata requirements (three parameters fully satisfied and one almost satisfied) already at the second stage of evolution and from the third step on, four requirements are fully satisfied. Concerning the other strategies that have an evolution with a specific goal, those that satisfy most of the requirements are the assortative high degree and the least distance one. Both strategies satisfy three out of the six parameters and the difference being that the assortative one has good performance for the characteristic path length parameters, while the least distance strategy reaches the target for the robustness aspects. Therefore, these sub-optimal strategies could be used where such different requirements are most needed.

From the comparison of the different evolution strategies and their steps in adding new edges, one notices that there are evolution strategies that tend to satisfy the majority of the metrics for the smart grid that we have defined (cf. Section 4.1). The evolution strategy that already in the second step fully satisfies three requirements is the random evolution. The requirements satisfied become four when 75% of additional edges are added. This strategy is in line with the finding of Section 4.2 where a small-world model was the best solution for modeling a smart grid topology. The addition of random edges goes into that direction: the network with the rational structure planned by the power engineers is modified by the addition of random links, thus in something between the rational topology and the fully random network. The other strategies that are quite successful are the assortative high degree and the least distance, but only when the edges are doubled. This is the situation considering a pure topological decision. When also the costs of the evolution are taken into the picture the optimal evolution strategy might change due to economic constraints. The extended analysis that considers economic aspects is performed in Section 4.3.3.

Considering the low voltage grid, the random evolution strategy is the one that satisfies most of the desiderata requirements at the final stage of evolution (two parameters fully satisfied and three almost satisfied). Concerning the other strategies that have an evolution with a specific goal, those that satisfy most of the requirements are the assortative low degree followed by the dissortative one which almost satisfy four and three requirements respectively. Therefore these sub-optimal strategies could be used too. In particular, the assortative low degree satisfies the requirements regarding the clustering coefficient, that are not satisfied neither by the random, nor by the dissortative strategy.

A commonality between the evolution for medium voltage and low voltage is the best result achieved by the random evolution of network that scores best among the strategies considered in this study. An interesting difference lies in the sub-optimal strategies that score high for the medium voltage and low voltage network evolutions. For the medium voltage network, it is best to provide more connectivity between those nodes that already hold a high node degree, thus reinforcing their role as key components of the network. On the other hand, for the low voltage network, it seems that adding more connectivity between the nodes that have a small connectivity provides better performances. The assortative low degree aims at creating more connections and (therefore hubs) where is less connectivity in the current samples. Also the other strategy that is sub-optimal for low voltage networks, the dissortative, aims at giving a more important role to the nodes in the periphery of the network by connecting them to the more connected nodes. A synthesis of the topological performance of the different strategies is provided in Table 4.21. For the layer of the power grid considered (column one), each strategy (column two) is assessed with an optimality level (column three) based on the satisfaction of the topological metrics. The step during the evolution process in which the most of the metrics are satisfied is also provided in column four. The achievement of good results depends on the network layer and on the strategy used. The optimality level is assigned based on the full satisfaction of the topological metrics described in Section 4.1. One can see a distinction between the medium voltage and the low voltage networks: the low voltage samples achieve less topological optimality (less metrics are satisfied) although requiring on average more evolution steps. Another aspect to note is that some evolution strategies (i.e., assortative high degree and triangle closure) do not have an improvement between the various steps despite the addition of more connectivity. On the other hand, other strategies (i.e., dissortative and random) benefit more from additional connectivity by improving the satisfaction of the metrics.

Network level	**Strategy**	**Optimality**	**Optimality reached at step**
MV	Assortative HD	⋆ ⋆ ⋆	4
MV	Assortative LD	⋆⋆	2
MV	Triangle closure	⋆⋆	1
MV	Dissortative	⋆	1
MV	Least distance	⋆ ⋆ ⋆	4
MV	Random	⋆ ⋆ ⋆⋆	3
LV	Assortative HD	⋆⋆	3
LV	Assortative LD	⋆⋆	4
LV	Triangle closure	⋆⋆	3
LV	Dissortative	⋆ ⋆ ⋆	4
LV	Random	⋆ ⋆ ⋆⋆	4

Table 4.21: Parameter optimal satisfaction.

4.3.3 Economic considerations

We need to characterize the cables used for the upgrade process in terms of their physical properties and costs, as we did in Section 4.2.3. The information of length for each cable in the existing sample networks is available. For the majority of the nodes in the samples belonging to the medium voltage we know the geographical coordinate information, so with some approximation it is possible to compute the length of the new cable connecting nodes that do not have a line connecting them yet. Therefore, each new added line is assigned a length. In order to define the properties characterizing the new cables, we adopt a k-nearest neighbor (KNN) classification based on the length of the cable [52]. The classification is performed using the Java Machine Learning Library (Java-ML) v0.1.5[3]. For each new cable, we consider the *m* cables in the original network used as the training set for the classifier that have the length closer to the added one and then the type of cable is provided by the classification algorithm. We choose this method to identify the cable since it

[3] `http://java-ml.sourceforge.net/`

is simple and it has proven successful in several practical contexts [189]. As the k parameter we choose $k = 5$ being a compromise towards speed of the algorithm execution in the trade-off between precision of the result and speed to achieve it. Once the properties of the new cables are identified, all the information we need for the economic analysis are then available: resistance per unit of length, cost per unit of length, maximal supported current. We remark that our proposed analysis does not aim at being a comprehensive investment analysis, for which other techniques are well established, simply an economic evaluation of the proposed evolution strategies to confirm their feasibility or unfeasibility.

To assess these costs in the medium and low voltage infrastructure, we consider the simple relation of Equation 4.1. As stated before, in this work we provide an initial estimate and simply consider cabling costs and ignore substation ones. In addition, the economic analysis can be applied only to the medium voltage network samples since there is no geographical information about the location of the nodes in the low voltage networks. Thus it is not possible to characterize length (and the associated cost) of new cables to be added in this layer of the network.

For the cost analysis, in the medium voltage case we limit our investigation to the three strategies that have scored best in the pure topological analysis (i.e., random, assortative high node degree and least distance). The results of the cost for the evolution of the Dutch samples are shown in Tables 4.22, 4.23, and 4.24. The first column of each table contains the sample ID, while the second provides the information about the evolution step considered, the third column has the information concerning the cost of the evolution of the network according to the specified strategy. The fourth column contains the information on the fraction of the cost that the evolution of the infrastructure impacts on the whole cost of the infrastructure. The most interesting result is the difference in the cost that the three methods of evolution require. The costs of network improvement is similar when considering the assortative and the random strategy: the cost of adding more edges is on average about 75% of the cost for cabling of the whole grid infrastructure. The situation is radically different for the least distance strategy whose development impacts only marginally in the total cost of the infrastructure, by adding just less than 13% of the cost of the infrastructure. Purely from the point of view of cabling costs, the most promising evolution strategy in economic and topological terms for the medium voltage grid is the strategy that connects the nodes that are geographically closer (i.e., least distance strategy).

One may then wonder if such investments are beneficial for the end users and the distribution companies in reducing the cost of electricity flows. We resort to a set of metrics that we have developed and already applied to the Northern Netherlands distribution grid and to synthetically generated networks. The goal is to consider

	Evolution step	Evolution cost (euro)	Fraction of evolution cost on the whole infrastructure
Sample #1	+25%	59334032	0.73
	+50%	59334032	0.73
	+75%	59334032	0.73
	+100%	119505699	0.85
Sample #2	+25%	107698129	0.73
	+50%	225049864	0.85
	+75%	255361286	0.87
	+100%	298792505	0.88
Sample #3	+25%	23007544	0.72
	+50%	44739934	0.83
	+75%	72515143	0.89
	+100%	105567690	0.92
Sample #4	+25%	45762304	0.65
	+50%	55159897	0.69
	+75%	55159897	0.69
	+100%	55159897	0.69
Sample #5	+25%	65393631	0.74
	+50%	117984090	0.84
	+75%	157876214	0.87
	+100%	194838593	0.90
Sample #6	+25%	31063470	0.67
	+50%	62971658	0.81
	+75%	104711687	0.87
	+100%	165938342	0.92
Sample #7	+25%	79148152	0.61
	+50%	199600692	0.80
	+75%	229365781	0.82
	+100%	350782953	0.87
Sample #8	+25%	32826377	0.67
	+50%	36976365	0.70
	+75%	83305090	0.84
	+100%	123241632	0.88
Sample #9	+25%	11739091	0.46
	+50%	42894090	0.76
	+75%	72834999	0.84
	+100%	103147797	0.88
Sample #10	+25%	14899834	0.60
	+50%	32136672	0.77
	+75%	45646927	0.82
	+100%	60546613	0.86
Sample #11	+25%	13094364	0.50
	+50%	26377193	0.67
	+75%	38709007	0.75
	+100%	40774034	0.76
Sample #12	+25%	93369553	0.70
	+50%	93369553	0.70
	+75%	138091611	0.77
	+100%	253056714	0.86

Table 4.22: Cost of medium voltage network evolution for assortative high degree strategy.

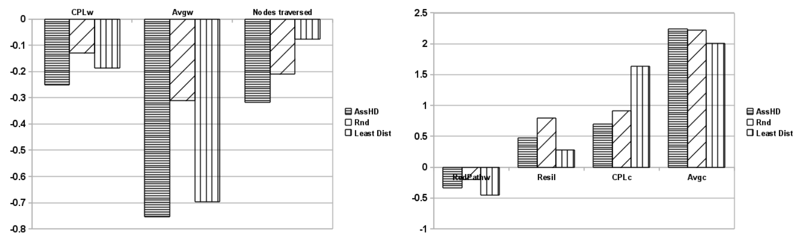
	Evolution step	Evolution cost (euro)	Fraction of evolution cost on the whole infrastructure
Sample #1	+25%	49585634	0.70
	+50%	99550100	0.82
	+75%	147237698	0.87
	+100%	193060613	0.90
Sample #2	+25%	81443232	0.67
	+50%	173835139	0.81
	+75%	270738228	0.87
	+100%	352561226	0.90
Sample #3	+25%	20401026	0.70
	+50%	39401799	0.82
	+75%	58698139	0.87
	+100%	80038042	0.90
Sample #4	+25%	38550574	0.61
	+50%	74405969	0.75
	+75%	112507292	0.82
	+100%	144092688	0.85
Sample #5	+25%	34665986	0.60
	+50%	63522559	0.74
	+75%	90171964	0.80
	+100%	118746816	0.84
Sample #6	+25%	32756917	0.68
	+50%	67956839	0.82
	+75%	102189138	0.87
	+100%	140632238	0.90
Sample #7	+25%	85124154	0.63
	+50%	181095750	0.78
	+75%	262020773	0.84
	+100%	352245673	0.87
Sample #8	+25%	40434688	0.72
	+50%	79435946	0.83
	+75%	107843650	0.87
	+100%	139828749	0.90
Sample #9	+25%	31096020	0.70
	+50%	57036433	0.81
	+75%	85190089	0.86
	+100%	115442294	0.89
Sample #10	+25%	12677249	0.56
	+50%	26126828	0.73
	+75%	42325971	0.81
	+100%	53357492	0.84
Sample #11	+25%	1214608	0.08
	+50%	2553383	0.16
	+75%	3117824	0.19
	+100%	4494600	0.25
Sample #12	+25%	105240341	0.72
	+50%	223059810	0.85
	+75%	317337397	0.89
	+100%	417627345	0.91

Table 4.23: Cost of medium voltage network evolution for random strategy.

	Evolution step	Evolution cost (euro)	Fraction of evolution cost on the whole infrastructure
Sample #1	+25%	635014	0.03
	+50%	1944353	0.08
	+75%	3536060	0.14
	+100%	5539077	0.20
Sample #2	+25%	1832132	0.04
	+50%	5362070	0.12
	+75%	9454081	0.19
	+100%	15488444	0.28
Sample #3	+25%	357093	0.04
	+50%	1143678	0.11
	+75%	2075073	0.19
	+100%	3206160	0.26
Sample #4	+25%	831466	0.03
	+50%	2936596	0.11
	+75%	5026907	0.17
	+100%	7688212	0.24
Sample #5	+25%	801764	0.03
	+50%	2128363	0.09
	+75%	3412290	0.13
	+100%	5187261	0.19
Sample #6	+25%	543105	0.03
	+50%	1433992	0.09
	+75%	2588723	0.15
	+100%	3822329	0.20
Sample #7	+25%	1104605	0.02
	+50%	3521507	0.07
	+75%	6534753	0.11
	+100%	10042275	0.17
Sample #8	+25%	584039	0.03
	+50%	2127943	0.12
	+75%	4016394	0.20
	+100%	5840352	0.27
Sample #9	+25%	661698	0.05
	+50%	1861427	0.12
	+75%	3476464	0.20
	+100%	5524643	0.29
Sample #10	+25%	314451	0.03
	+50%	987993	0.09
	+75%	1971402	0.17
	+100%	2912718	0.23
Sample #11	+25%	478534	0.03
	+50%	1230921	0.09
	+75%	2224697	0.14
	+100%	3262358	0.20
Sample #12	+25%	1069466	0.03
	+50%	3143898	0.07
	+75%	5723602	0.12
	+100%	8686081	0.18

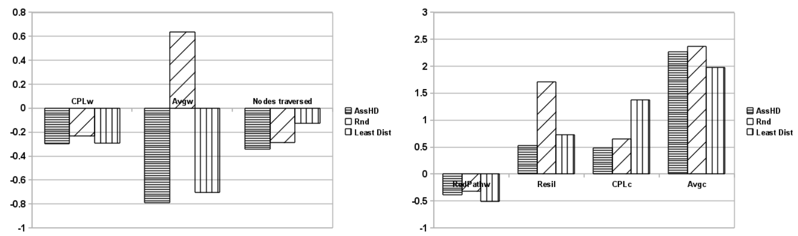
Table 4.24: Cost of medium voltage network evolution for least distance strategy.

those measures that are critical in contributing to the cost of electricity as elements in the transmission and distribution networks.



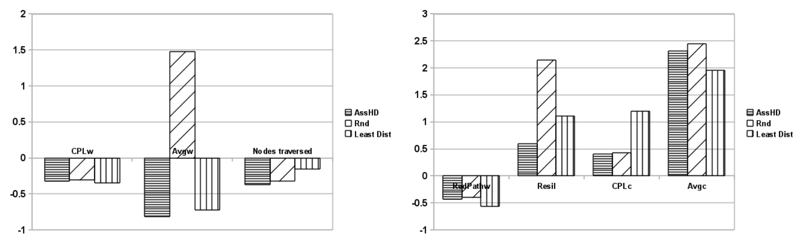
(a) Losses-related metrics. (b) Reliability-related metrics.

Figure 4.11: Comparison to the original networks (fraction of the original) of properties influencing electricity cost at 1st stage of evolution (i.e., +25% edges).



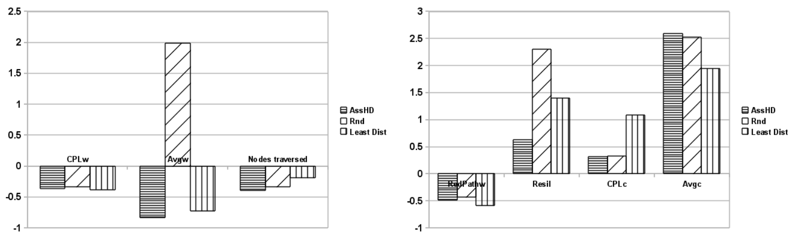
(a) Losses-related metrics. (b) Reliability-related metrics.

Figure 4.12: Comparison to the original networks (fraction of the original) of properties influencing electricity cost at 2nd stage of evolution (i.e., +50% edges).



(a) Losses-related metrics. (b) Reliability-related metrics.

Figure 4.13: Comparison to the original networks (fraction of the original) of properties influencing electricity cost at 3rd stage of evolution (i.e., +75% edges).



(a) Losses-related metrics. (b) Reliability-related metrics.

Figure 4.14: Comparison to the original networks (fraction of the original) of properties influencing electricity cost at 4th stage of evolution (i.e., +100% edges).

Figures 4.11, 4.12, 4.13, and 4.14 show the improvement as a fraction of the original values in the metrics related to the cost of electricity distribution respectively in the four step of evolution we consider (+25%, +50%, +75%, and +100% of edges) compared to the original samples. To give a general idea of the improvement, we use the average results over the 12 medium voltage networks samples. For each strategy, we look at the improvement, in percentage, for the same metrics compared to the initial samples. Already from a small increase in connectivity (addition of 25% of edges) the resistance of a path decreases with every strategy with the best result obtained by the assortative high degree strategy. Adding more lines in general promotes a reduction of the average resistance of edges compared to the initial situation. This is even more true in the assortative and least distance strategy where short distance connections (and therefore less resistive) tend to dominate, or more efficient cables are chosen in the KNN procedure selection. In addition, a considerable reduction in the losses that are experienced in traversing substations can be avoided with more dense networks: almost a 30% less station traversed in the graph enhanced with assortative connections. Of course less benefits take place with the least distance strategy that reduces the average number of traversed stations in a path by 7%. This higher connectivity provides benefits to the resilience and robustness aspects: resilient paths (i.e., 10th shortest weighted path) that are less lossy (about 30% reduction for the least distance strategy). An improvement is also experienced in the resilience of the network to node disruptions: about 80% more robust with the random strategy. Also the amount of current that is supported in a path (the weighted characteristic path length with maximal supported current as weight of the edges) is definitely higher (more than 2.5 times for the least distance strategy) compared to the initial samples.

Similar considerations can be done for the second evolution step (i.e., +50% cables). Considering the loss aspect the benefits are for the three strategies on average about 30% in the reduction of the weighted path. For the average weight (i.e., resistance) of a cable the random addition creates networks with much more cables with an high resistance. This is an indication that the random strategy creates long distance connections that act as shortcuts in the network from a topological point of view. However, in a physical system there are no significant benefits since the resistance grows with the distance. These limits of the benefits are shown by the weighted path analysis. Considering reliability-oriented metrics, the most of the benefits take place in resilience where the networks on average stay twice as much as connected than the initial evolution step, except for the assortative strategy that lags behind. We see a small decrease for all the strategies in the characteristic path when the weight is the maximal current that can flow in the cables; this is not difficult to understand since the characteristic path uses the most efficient (i.e., with smallest weight paths).

When even more connectivity (+75% and +100%) is added one sees even better results and improvements but smaller in their magnitude. In these last two stages, the benefits for the losses reach values of reduction above 30%, the same for the nodes to be traversed on average in a path. Only the least distance strategy has a small reduction, but this is due to the very nature of such evolution strategy that avoids long distance (topological shortcuts) cables. Considering network robustness, the main improvements are for reliability to node failures that for the least distance strategy allow the network to be twice as much as robust than the initial samples. Concerning the capacity of the network, we see a slight decrease for the same reasons as above. However, the networks with the least distance evolution are able, on average, to transport twice the amount of current than the initial samples.

We can now provide a recommendation on how to evolve the analyzed samples given the cost analysis for cable addition and the benefits that such an higher connectivity brings in the topological aspects related to the price of electricity. Using our topological-based method, we consider that the least distance strategy with the addition of 75% or 100% of edges is a very good way of making the network more connected. Such an evolution strategy provides benefits from a topological point of view, and it keeps the costs low. For the samples examined, it requires an investment in cabling costs about 25% of the value of the cables already on the ground. Such investment can provide benefits in the loss reduction of the network as well as in its robustness. The energy economic studies show [84, 141] that losses and reliability are directly related to the cost of electricity. These factors are tight to topological parameters as shown above, therefore such evolution would provide less cost in electricity distribution. It is difficult to translate in exact monetary terms for the end user concerning the savings in the energy bill that an improvement in the topology parameters that influence the cost of energy distribution might bring. For the realization of a smart grid less costs in the distribution of electricity and a local reliable and robust network are the essential ingredients to enable a paradigm where energy is produced and distributed locally such as at neighborhood or city level.

In this analysis of the influence of the topology on the price of electricity, we have not used the α and β metrics themselves (cf. Section 3.4), but the components of these metrics. In fact, especially for the α metric we see a decrease of the average weight of the edges compared to the Dutch grid samples for when the evolution is performed (exception is the random strategy where long, therefore high-weight connections are present). The decrease in the average weight of edges is faster than the decrease in the average path length, therefore the L_{line_N} part of the α metric grows, which is counterintuitive given the decrease that all the components characterizing the losses show. Therefore, we have found more appropriate and meaningful for the analysis of the evolution to consider the single components influencing losses and reliability and not their aggregation.

4.4 New Topologies for the Smart Grid

In an evolving electricity sector with end users able to produce their own energy and sell it on a local-scale market, the grid plays the essential enabling role of supporting infrastructure. Local-scale energy exchange is in fact beneficial for several aspects such as the increase in renewable-based energy production, the possibility for the end user to have an economic contentment by selling surplus energy and, not less important, a step forward to the unbundling of the electricity sector. We studied how different topologies inspired from technological and social network studies have different properties and can be (or not) suitable for the future smart grid networks. We showed that between the various models analyzed, the small-world model appears to have many supporting characteristics, according to a set of topological metrics defined for power grids. We also showed how these topological benefits can be related to economical aspects of electricity distribution.

We have also shown that different strategies to evolve the current physical grid provide different cost/benefits ratios and have identified which strategies are mostly promising. The statistical approach that we take, combined with traditional power engineering approaches, can form the basis of a decision support system for evolving the current power grid into a “smart grid friendly” network.

We have seen that the random and assortative high degree evolution strategies are the best in a topological weighted analysis, but they have costs that are extremely high and unrealistic to realize in practice. In addition, the weighted analysis gave us the possibility to investigate the elements of the power grid that influence the cost of electricity distribution (i.e., grid losses and grid reliability) and the results of this analysis suggest that evolving the network by adding connections between the nodes with smallest distance is beneficial and provides in many cases better results compared to the other evolution strategies. Therefore, with an investment about 25% of the actual costs in cables already on the ground the distribution grid can improve consistently in reducing transportation costs.

From the industrial perspective, where a unique and clear definition of the term smart grid [139] is missing and where the standardization process is at the early stages of development, we consider that the present proposal is usful to make general decisions on how to evolve the grid and what costs are entailed at least from a coarse grain point of view. Existing planning techniques will have to be revised in the future, especially for the distribution grid due to the presence of Advanced Metering Infrastructure (AMI) (i.e., bidirectional intelligent digital meters at customer location) and Distribution Automation (DA) (i.e., feeders can be monitored, controlled in automated way through two-way communication). In addition, the medium and low voltage grid will no longer be a layer where only energy is consumed, but distributed energy generation facilities (small-scale photovoltaic systems and small-wind turbines) will be attached to this segment of the grid; altogether these elements are likely to reshape the way planning for medium and low voltage is realized [31] and will also call for new instruments such as the one we propose here.

Chapter 5

ICT Services and Applications for the Smart Grid

One of the main innovation of the smart grid will be to bring more Information and Communication Technology (ICT) capabilities and computation in the power domain, especially in the low layers of the grid and at the consumer facility.

We consider four different aspects to show how the ICT is important for the success of the smart grid and in which scenarios it is a key component. We consider the benefits and necessity of a software approach based on service-oriented architectures (SOAs) in Section 5.1. We describe in Section 5.2 an application of services for the realization of demand-response functionalities and an implementation with today's technologies and actual data services. We further describe how these smart grid services can be used in a real office environment to realize a demand-response solution with the goal of minimizing the energy expenditure while keeping users' satisfaction (Section 5.3). We conclude the chapter with another application of the smart grid such as an energy market where users provision energy through agent technologies and automated ICT platforms.

5.1 The Services: the Future Energy Landscape

The vision of the future smart grid, however defined, will bring unbundling of energy markets, appearance of renewable generating facilities at all scale levels, the standardization of the control elements of the power grid, the diffusion of digital/data prone smart meters, and especially two way communication among end users and the grid. This will be a reality only if there are infrastructures that can support it especially from a communication and information point of view.

5.1.1 Smart grid challenges

A number of challenges needs to be addressed for building such a grid. We identify the following ones as the key issues to be solved at the software level.

- **Interoperation.** The number of actors populating the energy market is constantly increasing and their capabilities as well. There is a strong need to have standards for interoperation at all levels (not only the control layer of the grid). Furthermore, standards tend to cover the syntactic part of the interoperation, while the semantics of the message exchange is scarcely addressed.

- **Scalability.** The increase of actors also involves scalability issues. If millions of micro energy producers start trading micro-quantities of energy, there must be an appropriate infrastructure to manage this, possibly real-time, information exchange.

- **Discovery.** If the actors increase and more entities can take on the same role, one may think of discovering services on the fly. The idea of signing an yearly contract for a home, may be too limiting and one may want to switch energy supplier on a much shorter time frame. Furthermore, if anybody can be a supplier, then one may want to find a provider in the moment the energy is needed.

- **Mobility.** In the future, grid energy consumers, and also producers, may be mobile on the grid. Cars will be electric, but may also have energy producing and storing facilities (e.g., a solar cell roof, fuel cells powered engine). The mobile elements need to interact with the power grid in a transparent way.

- **Resilience to failure and trust.** The electrical power grid is a critical infrastructure. A key performance indicator of the current energy distributors is the down time that should not exceed the few hours per year. When moving to an open smart grid the delivery of energy must not decrease in quality. This requires having a trust mechanism among the various players. It may also require having reliable forecasting of generation and use.

- **Service integration and composition.** The physical layer, the data layer and the business layer will have to interact more closely. In fact, any node that produces energy needs to interact with control/actuation part of the grid and get paid for the energy produced; the generation might be a part of a larger business process relying on the energy (e.g., one could drive an electric car while lodging at a motel, plug it into the grid and use the car generated electricity to pay part of the bill [208]).

- **Topology.** The current infrastructure is highly hierarchical, not only in the physical infrastructure, but also in the information systems that manage the electricity system. Very few large energy producers and backbone owners exist, with few systems that control the plants and the transmission network is highly centralized. But the new vision of the open grid demands for a flat peer-to-peer (P2P) network in which all actors are producers and consumers of energy, data and services.

- **Smart Meters.** The smart meter is likely to become the "energy gateway" of the house with more and more functionalities and embedded intelligence. Smart Meters might work as automatic bidder on the energy market knowing family energy usage and production patterns together with estimation methods based on past usage and environmental forecasts (e.g., future weather conditions).

- **Real-time.** Energy related operation such as control, actuation, distribution and trading have very strict time-dependent constraints to satisfy. All the players in the next generation grid must interact following real-time commitments to provide and receive an energy service with the proper quality.

5.1.2 Service-oriented architecture supporting the smart grid

Interestingly, the just listed challenges have a natural counterpart in the service-oriented architectures. These have been traditionally built to address interoperability and scalability issues for the integration problem of enterprise information systems (e.g., [173, 41, 172, 174]) or to support business process, especially across companies borders. Here we take a different look and consider how SOAs are appropriate for the smart grid and look at the challenges just introduced through the glasses of service-orientation.

- **Interoperation.** Web services are a technology to build service-oriented architectures and address the problem of interoperation being standardized eXtensible Markup Language (XML) protocols to describe messages, remote operations and coordination among loosely coupled entities, e.g., [120]. These are already entering the energy sector as described in Chapter 2.

- **Scalability.** The basic SOA pattern: publish–find–bind allows to decouple service consumers from producers and to substitute, even at run-time, one component for another one. The communication, most often asynchronous, provides all the ingredients for a highly scalable infrastructure. Examples of which have already appeared in the area of eBusiness.

- **Discovery.** Discovery is one of the basic ingredients of a SOA. It needs to support the publish and find operations and is usually based on registries, but can also be realized with flooding models.

- **Mobility.** A SOA supports actor loosely coupling and behavior based binding, therefore the mobility of the elements is easily supported, e.g., [1].
- **Resilience to failure and trust.** Protocols exist to enable a Web service based SOA with trust, privacy and security support. This can provide the basic for a secure infrastructure. Reliability will also have to be pursued with appropriate energy technology which is beyond the SOA.
- **Service integration and composition.** Service integration and composition is the key added value of a SOA and many examples exist on methodologies to support this, e.g., [63, 2, 114, 20, 37].
- **Topology.** SOAs support any kind of topology. The hierarchical client-server one is less common, but can be realized. The P2P topology is most often the one realized.
- **Smart Meters.** In the SOA paradigm the smart meter is basically a service provider and a service consumer at the same time. It invokes other services to interact on the market and also provides services to other market participants interested in energy purchase. It also interacts in a service-oriented fashion with intelligent home appliances that require energy at a certain time.
- **Real-time.** Solutions are available to introduce enhancements to SOA paradigm in order to provide an appropriate quality of service and satisfy real-time constraints, e.g., [205, 171].

5.1.3 Traditional service-oriented architecture vs. energy service-oriented architecture

The smart grid is thus amenable to be supported by SOAs, though there are some differences with traditional SOA approaches. Table 5.1 shows the main points of contact and dissimilarity between the traditional SOAs and those for the energy sector. A natural common point is the use of SOA technologies for integrating heterogeneous systems thus enabling their interoperability. Beyond this common feature, several differences lay that must be taken into account when dealing with energy systems. Traditional SOA is mainly used in the business process domain managing complex supply chain and interactions between a multiplicity of actors whose applications are usually triggered by specific events. Usually, the paradigm of these interactions is asynchronous. On the other hand, the SOA for energy applications must tackle some peculiarities of this type of business and systems. First of all the requirements for real-time interactions between the various subsystems and components of the energy-related ICT involve Supervisory Control and Data Acquisition

(SCADA) and Energy Management System (EMS) systems and low level electric applications embedded systems. Being these systems highly important and mission critical, real-time constraints together with fault tolerance, security and trust mechanisms in the service provisioning are essential requirements that a SOA for the electricity sector needs to satisfy. An interaction with non-strictly related energy systems such as Customer Relationship Management (CRM) and Enterprise Resource Planning (ERP) is also required to have a complete interoperability picture.

Another issue that is central in enabling the SOA solutions for the smart grid is of course an appropriate communication infrastructure. It is not the focus of the present work, but it is worth to mentioning the adaptation required by the telecommunication/telecontrol infrastructure to support the enhanced amount of information data and control signaling that the smart grid requires [227]. The telecommunication aspects and its infrastructure must not be taken for granted, since they form the basis to build more complex service-oriented software layers on top.

Another aspect that is likely to appear in the new smart grid landscape is the increase in the interactions that follow a P2P paradigm. In fact in a grid with many more energy prosumers and the related information exchange, end users will be directly involved in a fashion similar to the P2P paradigm in data exchange. Although the high level picture might look the same, some important differences remain beyond the main commonality that is the scalability requirement. Traditional P2P architectures are used in the framework of data exchange with requirements in the infrastructure that lay between the static and dynamic solutions based on the specific application. This aspect usually reflects on the network structure that the P2P system acquires (i.e., structured, unstructured). The energy P2P infrastructure has to deal with energy flows which imply, at least with actual technologies, the extreme difficulty in energy storage and the seek of an energy balance equilibrium all the time; on the other hand when dealing with P2P for information exchange, data buffering and replication are easy and cheap options to increase system performance. The dynamism in the energy paradigm is high when considering the demand side, but it is almost static considering the locations of the user. The result is a static network structure (an exception lies in the extensive usage of electric vehicles that add dynamism also at the location aspects). Performance and security are two key aspects that differentiate the data and energy P2P worlds: these properties are tight to the specific application considered for the former, while for the latter real-time requirements and high security are essential to deliver a valuable energy service. The level of centralization in a P2P data environment is closely coupled with the application it has to satisfy; in the energy environment the system tends to evolve to an hybrid solution since generating companies that own power plants still

<table>
<tr><th colspan="2">Traditional SOA Vs. Energy SOA</th></tr>
<tr><td colspan="2">Similarities</td></tr>
<tr><td colspan="2">• System integration
• Interoperation</td></tr>
<tr><td colspan="2">Differences</td></tr>
<tr><td>• Supply Chain management

• Event-based applications
• Business process management

• Asynchronous long-running business process and transactions</td><td>• Interaction with non-Energy systems (e.g., ERP, CRM)
• Real-time requirement
• High security (authentication, encryption for market trading)
• Trust mechanism for services provided

• Interaction with low level electrical interface standards (IEC standards for SCADA/EMS)
• Fault tolerance</td></tr>
</table>

Table 5.1: Similarities and differences between Traditional SOA and Energy-Oriented SOA

act as sort of super-peers in the system.

5.2 Today's Services for Implementing the Smart Grid

The smart grid and the SOA-based vision described in Section 5.1 are far from being implemented. At most there are local and small scale testbeds (e.g., Hoogkerk's PowerMatching city smart grid project [99]) where some features of the smart grid are implemented. In addition, business analysis and business cases around the smart grid are still under investigation [29, 147]; so is the standardization process. The smart grid is still rich of uncertainties, however some functionalities are assumed to be its pillar: demand-response through dynamic pricing, local energy generation through renewables, and energy forecast to help schedule smart appliances when more favorable. In such scenario, our aim is to design and realize a Smart Grid Simulation Engine based on available Web resources. We have developed ad hoc interfaces to existing services on the Web (e.g., energy tariffs, meteorological forecast for renewable energy) in order to realize the smart grid features essential for a smart meter/smart home interaction. In particular, information regarding energy prices, energy generated, and environmental conditions is accessed through Web sites or Web services. For each smart grid functionality such as energy price, energy production metering, and meteorological conditions, we describe the solution we envision to be available in the future, but we also describe what can be realized today with the currently available information. A recurrent aspect for the envisioned ideal solution is the use of SOAs. A typical implementation of SOA is through Web services. An incarnation of Web services is through protocols like the Hypertext Transfer Protocol (HTTP), the Simple Object Access Protocol (SOAP), and through standardized service descriptions interfaces such as the Web Services Description Language (WSDL). A summary of the vision for the smart grid services, the standardization process involved, and the nowadays possible implementation is shown in Table 5.2.

Category	**Vision**	**Ongoing standardization effort**	**Nowadays realization**
Energy price	Third party Web service providing prices from several energy providers	OpenADR, ZigBee Smart Energy Profile 2.0, OASIS EMIX, IEC-62325, direct interaction with smart meter	Ad hoc data retrieval in real wholesale energy market
Energy production/ consumption	Web service in the control unit of the small-scale production plant	OpenADE, ZigBee Smart Energy Profile 2.0, IEEE 1547.3, direct interaction with smart meter	Ad hoc interaction with producing/consuming device control and monitoring equipment, interaction with smart meter of a specific standard/technology
Environmental conditions	Dedicated Web service from meteorological companies specialized in energy with very localized forecasting	Outcome of NIST B2G DEWG, addition to existing weather standards e.g., Weather Information Exchange Model (WXXM)	Interaction with existing weather Web services (e.g., Yahoo!, NOAA, The Weather Channel)

Table 5.2: Smart grid services: vision, ongoing standardization processes, and nowadays possible realization.

5.2.1 Interface towards energy price service

The availability of updated prices and price forecast within few hours or within few days are essential to implement the smart grid demand-response functionality.

Vision

The envisioned solution to implement the price mechanism would be a Web service provided by a third party for the smart grid. This party administers the energy tariffs for the providers that the smart home is allowed (or wants) to supply from. The smart home subscribes to this service indicating for which providers the tariffs are required. Either in a pull or push method the smart home obtains the forecast tariffs for different time granularities. The granularity can be real-time for balancing purposes or hourly day-ahead forecast for mid/long term appliances schedule. The service, in addition, signals through a push message a changing tariff compared to the previous forecast so that energy intensive operations already scheduled in the house can be adapted on-the-fly to the changing price conditions. In this ideal situation all the interactions and information exchange take place automatically thorough the World Wide Web and the smart meter. The OASIS Energy Market Information Exchange (eMIX) [154] might be the protocol used to exchange information about energy prices.

Realization today

The solution proposed above is ideal and not implemented at the moment. However, it is possible to have a similar behavior by interacting with wholesale energy markets that already implement price differentiation in the energy trading process. In particular, the prices generated on the wholesale market vary in accordance with the congestion and energy availability on the power grid infrastructure. This is realized to keep the balance between demand and supply, in other words the demand-response mechanism. To have variable energy tariffs it is possible, for instance, to use prices coming from the PJM Interconnection[1] which is a Regional Transmission Organization (RTO) that coordinates the movement of wholesale electricity in more than 13 states of Eastern U.S.A., or from another organization such as Independent System Operator New England (ISO-NE)[2] which provides the energy market facilities for the New England region of the U.S.A. For both markets the data that can be extracted are the real-time prices and the day-ahead energy market locational marginal pricing (LMP). These prices are respectively the live minute-by-minute prices and the prices of energy negotiated in the wholesale market for the following day by energy companies at specific locations where energy is delivered or received. For the day-ahead market, data contain the energy price for each unit (dollars per Megawatt-hour) for each hour of the day (for the next day) at the locations of delivery. Each location can have in principle a different price from any other one. These prices can be considered for the smart home as the prices of different energy providers. Real-time prices can be obtained by repetitively interrogating the PJM or ISO-NE Web site where prices for the various location in the network are continuously updated; whereas the day-ahead data can be automatically obtained through PJM or ISO-NE Web site and they are available each day for the following day (day-ahead) as a comma separated value (CSV) files. Little ad hoc adaptations are required to automatize the process of file download and cleaning of the unnecessary

[1] http://www.pjm.com/
[2] http://www.iso-ne.com/

fields, before they can be used as the input of a systems that provides dynamic pricing for demand-response purposes. A feasible solution today is to use day-ahead prices to forecast the schedule for smart appliances for the following day, and then correct the scheduling in real-time if considerable difference in prices arise in the real-time market.

5.2.2 Interface towards energy producing equipment service

Another essential feature for a smart home is the possibility of generating energy through small-scale renewable plants (e.g., photovoltaic (PV) panels, small wind-turbine).

Vision

The envisioned solution to implement information retrieval of the amount of energy and power produced by a small-scale energy plant would be a Web service provided by the control unit of the plant. The smart home, directly or through the smart meter, could then in an automated way ask the information about the available instantaneous power and the amount of energy generated. Information about the generated power could be easily and automatically reported via Web service interaction to the utility company with whom the sell contract is signed if no Advanced Metering Infrastructure (AMI) system is already implemented.

Realization today

Usually, every modern small-scale plant has an electronic controller that is able to show information about power and energy produced together with essential technical and environmental parameters of the plant (e.g., for a photovoltaic plant: air temperature, panel temperature, and solar irradiance). This information is usually published on the Internet (or LAN) through the Web server embedded in the controller of the plant so that the owner of the plant can monitor the performance. This information can usually be directly accessed or downloaded in CSV format so with little ad hoc manipulations these data can be used as input for an application. An example is available at the U.S. School Power Naturally data portal.[3] Through this portal it is possible to access solar plant information and data. Another example of a Web server is available at `http://pagani.dyndns.org/html/en/onlineOverWr.html`, it provides real-time and historical information from a solar plant about the power and energy produced. While a comprehensive small-scale power generating

[3] `http://sunviewer.net/portals/NYSERDA/`

unit composed by a solar, wind turbine, and small-scale combined heat and power turbine plant is available at Grand Valley State University.[4]

5.2.3 Interface towards environmental service

External environmental and weather parameters are essential to know how external conditions may influence both energy consumption inside the smart home and the renewable-based generating equipment. The smart home has then to be able to access current and forecast environmental conditions.

Vision

The envisioned solution to obtain environmental parameters would be a dedicated Web service provided by a third-party specialized in environmental conditions forecast for the smart grid with precise geolocalization. In addition to traditional weather information (e.g., weather conditions, temperature, pressure), other information such as solar irradiance, wind speed, humidity, cloud coverage, air density, external light conditions and others are significant to estimate energy consumption of buildings and to forecast energy production by small-scale energy equipment. This information should be extremely localized and dedicated to the specific location of a smart home. The envisioned service requires that the smart home subscribes to the service by providing exact location (longitude and latitude coordinates) and orientation towards the sun. Once subscribed, the service should be able to provide the environmental parameters mentioned above. The service could be implemented in a push and pull manner so that updates on the conditions and forecasts can be sent to the smart home regularly, or it can request for an update when needed.

Realization today

On the World Wide Web there are several weather services, with advanced pieces of functionality that can be accessed in the form of Web services. A primary example, considering the levels of detail and abundance of weather-related parameters (current weather, 7 day forecast and many parameters such as max and min temperatures, wind speed, wind direction, wind gust, percentage of sky cover and more), is provided by the National Weather Service realized by the National Oceanic and Atmosphere Administration (NOAA)[5] in the U.S.A. From a technical point of view the service provides an interaction through SOAP requests. Other services that use

[4] `http://datamonitoring.marec.gvsu.edu/`
[5] `http://www.noaa.gov/`

a Web service approach are provided by commercial services such as Yahoo!.[6] This weather-related information seems the closest to the ideal solution for this kind of service required by the smart home in its interaction with the smart grid even if the precise localization is not realized yet.

5.2.4 Implementation

Proof-of-concepts implementations of the smart grid, simulations of customer behavior or small scale pilot projects have been reported being underway [124, 203, 110], though no actual large implementation exists.

To be able to perform research and development related to smart grid, we have developed smart grid services considering the case of a smart home located in New York, U.S.A. with the ability of generating energy with small-scale production units through a photovoltaic installation of 2.4 kW of power realized using an AstroPower AP-100 PV module,[7] and a small-scale wind turbine namely the Proven 2.5 wind turbine.[8]

In our simple implementation scenario for smart grid interaction we decided to use data and services coming from real markets, real energy installations, and real weather information. Such choice enables us to realize a realistic simulation of dynamic pricing functionality, energy production from small-scale renewable sources, and energy forecast. A representation through a block diagram of the components involved and implemented in software using JAVA programming language is shown in Figure 5.1. The central component is the Smart Grid Simulation Engine which is responsible for contacting the information providers. Its role is to establish external connections, gather the data and convert them to the format required by the internal sub-component that has specific tasks to further process those data.

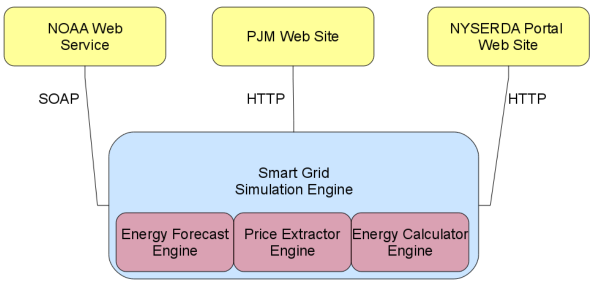


Figure 5.1: Smart Grid Simulation Engine and external services.

[6] `http://developer.yahoo.com/weather/`
[7] `http://atlantasolar.com/pdf/Astropower/ap-100.pdf`
[8] `http://www.windandsun.co.uk/Wind/wind_proven.htm`

To simulate the variable energy tariffs, we use the energy prices coming from the PJM Interconnection and the data extracted is the Day-Ahead Energy Market LMP. The Smart Grid Simulation Engine makes an HTTP request to the PJM Web site where day-ahead LMP prices are stored as CSV files. The data contain the energy price for each unit (dollars per Megawatt-hour) for each hour of the day (agreed for the following day) at 20 locations of delivery. We consider that the energy supply of a home could be fulfilled not just by one provider at a time, but also by the composition of several providers that might contribute to satisfy the power required by a house (e.g., provider A is able to give 1 kW of power at tariff x and provider B is able to give 2 kW of power at tariff $x + \Delta$ for a certain time period).

To account for the energy produced in the smart home, we access a real PV installation in New York at Dalton School in Manhattan[9] that has the same PV array hypothesized in our smart home and whose real-time data can be accessed through the U.S. School Power Naturally data portal. The Smart Grid Simulation Engine connects to the Web site[10] providing the performance of the solar installation on top of the school, and extracts the CSV file with real-time information about the power generated by the PV array. Data is sampled every 15 minutes. Information about relevant environmental parameters such as solar irradiance, ambient temperature and wind speed is also available. For our purpose, we assume a simplified situation: the data about the power is available every hour and constant.

To simulate the power supplied by the wind turbine in the smart home, we use the data gathered by the anemometer at Dalton School. The CSV file with the data is obtained through an HTTP request issued by the Smart Grid Simulation Engine towards the school Web site. The computation of the power produced by a wind turbine, given the wind speed sensed, is performed by the Energy Calculator Engine component. The component uses the well-know relationship between wind speed and power extracted by a wind turbine:

$$P = \frac{1}{2}\rho A U^3 C_p \tag{5.1}$$

where ρ is the air density, A is the rotor swept area, U is the wind speed and C_p is the power coefficient representing the efficiency of the turbine rotor [72]. Once we have chosen the turbine, the parameters are known: $A = \pi(\frac{3.5}{2})^2$ (the turbine blades have a 3.5 meters diameter), $\rho = 1.225$ (typical air density value), $C_p = 0.35$ (a typical value of rotor efficiency for wind turbines) and the wind speeds (i.e., cut-in and cut-out speed) between which the turbine works, then the wind to power relationship can be applied. Also in this case we assume to have the data about the wind speed available every hour and constant.

[9] http://www.dalton.org/

[10] http://sunviewer.net/portals/NYSERDA/siteHome.php?sid=17

5.2.5 Pricing energy from renewables

Usually, renewable sources are considered as *base load* in energy dispatching mechanisms: when energy is generated by these types of plants it is fed into the system and these plants are never shut down [200]. In our smart home, we consider both solar and wind power to have a cost of production due to the investment that is required to purchase, install, and maintain the equipment.

The idea, and the related computations performed by the Price Extractor Engine consider a simplified investment analysis for calculating the cost for each kWh of produced energy. We simply consider the investment cost, the maintenance and operations costs, governmental incentives and the energy produced over the expected lifetime of the plant. Therefore, we have the energy cost coming from the PV as

$$EC_{PV} = (C_{Inv} + C_{Opr} + C_{Mai})/En_L \tag{5.2}$$

where C_{Inv} is the total investment costs for the PV array, C_{Opr} and C_{Mai} are the total cost of operation and the total maintenance costs respectively over the investment lifetime for the PV system. For the PV array we consider 20 years of service for the AP-100 model. En_L is the estimated overall energy to be produced during the lifetime of the PV array. First, we estimate a production of energy during the 20 years of panel lifetime that is on average the same as the one produced in the previous years since the installation at Dalton school. Second, the investment cost is based on the results by Wise *et al.* [225] that investigated the cost of PV panels in the U.S.A. The value that emerges from their analysis considering the cost for PV panels, inverters and installation once the incentives applied by the U.S. government are subtracted, is 5.1 dollar for each installed watt of power. Third, the maintenance and operation costs can be considered as an annual expense about 0.84% of the investment, as found by Rehman *et al.* [181] for a PV installation. Another component that might be associated to operation and maintenance costs is the replacement of inverters which, as suggested in Croxford *et al.* [55], should be considered every 10 years. According to the findings of Wiser *et al.* [225] inverters replacement amounts about 7% of the initial total investment in a PV project for residential and small commercial purposes. Once we feed these parameters in the Energy Calculator Engine component it provides a constant value that represents the cost of each kWh generated by the PV plant. Therefore, the PV plant can be considered as an additional "virtual" energy provider since the produced energy in a future deregulated smart grid could be sold on the market at the price of returning from the investment (i.e., obtained from the computation shown above) or higher to make a profit.

We apply a similar relationship to evaluate the cost of electricity generated by the small-wind turbine:

$$EC_{SW} = (C_{Inv} + C_{Opr} + C_{Mai} - I_{Gov})/En_L \quad (5.3)$$

the only parameter that we have in addition to those described above is I_{Gov} that represents the incentive subsidy from the government in wind energy. We consider an investment cost for the wind turbine of 28000$ comprising the installation and required equipment (e.g., installation pole, inverters, workforce) and an annual cost of maintenance and operations of 1% of the total investment. This is in line with the findings of [213]. En_L for the turbine is the estimated overall energy to be produced during its lifetime (according to [187] the lifetime of a wind turbine can be considered around 20 years). To estimate the production of energy in New York we consider the average wind speed recorded by NOAA in more than 50 years.[11] Using this data the turbine produces around 1300 kWh of energy per year. Considering the subsidy, the state of New York has an incentive program for small-scale wind installation to promote such energy source with a contribution for one year of 3.50$ per kWh produced.[12] Therefore, with all these parameters fed in the Energy Calculator Engine component one has once again a constant price for each kWh generated, in this case by the small-wind turbine.

5.2.6 Renewable energy forecast

The amount of energy that can be produced by renewable systems is strongly correlated to the meteorological conditions. It is beyond the scope of this work to build an exact model of how meteorological and environmental conditions influence the energy production of wind and solar plants; here we aim at a realistic estimation of possible energy production in the broader context of simulation of realistic services for the smart grid. The Smart Grid Simulation Engine interacts with the NOAA national weather service[13] Web service requesting information about the wind speed, cloud coverage and temperatures. The service provides a week forecast given the location of interest (through longitude and latitude coordinates); for the first three days the forecast data has a 3 hours interval, while for the following days the interval is 6 hours. The Energy Forecast Engine has the task to adjust the meteorological data to an hour-by-hour information (in our simple solution we use interpolation) and to compute an hour-by-hour energy forecast for the week. The available power for each hour is computed in the following way:

[11] http://lwf.ncdc.noaa.gov/oa/climate/online/ccd/avgwind.html
[12] http://www.dsireusa.org/incentives/incentive.cfm?Incentive_Code=NY35F&re=0&ee=0
[13] http://graphical.weather.gov/xml/

- for the wind turbine, we apply $P = \frac{1}{2}\rho A U^3 C_p$, given the wind speed from the forecast and the other wind turbine parameters;

- for solar panels, we apply $P = H_{pw21}[(1 - C_c) + T_f]$ where H_{pw21} is the hour-by-hour average of the historical power produced the previous year during the 10 previous and 10 following days of the very same day of the previous year. Therefore, we have an average value of produced power in comparable days, i.e., similar sun horizon condition. The terms in square brackets C_c and T_f are used to correct the power considering cloud coverage and temperature estimations. C_c represents the percentage of cloud coverage and $T_f = (T_{his} - T)\delta$ is a temperature factor that considers, based on the difference in the average hourly temperature in the same 21-day period in the previous year, the increase/decrease in efficiency of δ percentage. We consider a linear relation which establishes a decrease/increase of $\delta = 0.5\%/°C$ in PV efficiency for a polycrystalline silicon panel as the temperature increases/decreases compared to factory test conditions [106].

Following these models, the Energy Forecast Engine provides for one week available power hour-by-hour for the two sources of renewable energy installed in our hypothetical smart home.

5.2.7 Examples

Figure 5.2 shows an example of the output of the Price Extractor component of the Smart Grid Emulation Engine. It represents the price of electricity on the PJM market for a sub set of the zones in the U.S. electrical system. One notes that there are significant variations in price during the day: from the minimum in the middle of the night, the price almost doubles in the evening. Therefore, with such information available beforehand (we recall that these prices are day-ahead) one can create automation to provide consistent savings over time. Think for instance of the scheduling of a washing machine, a dryer, and, a dishwasher, or the charging of an electric car.

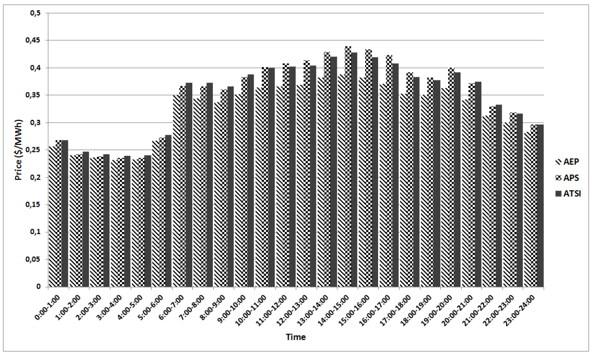


Figure 5.2: Energy prices for several locations of the U.S. electrical system on 5th October 2012.

Having information about the amount of renewable energy to be produced (forecast) is essential to better plan on a week basis the activities that can wait the optimal environmental conditions for self-production. Figure 5.3 shows the amount of energy that is forecast at the installation site we consider (Dalton school) and by using the weather forecast of NOAA and the model proposed in Section 5.2.6. In the figure, one week of power forecast is shown. The continuous line represents the power provided by the PV installation with a typical bell-shaped power output during the daytime and flat night profile. The dashed line represents the forecast of the wind turbine that has an irregular pattern. One can see how 6th, 7th and 10th October are more favorable to the self-production of energy by using solar panels, while the night between 6th and 7th and the days 10th and 11th are very favorable to power production through the wind turbine. By exploiting this information the smart home can have an idea about the right time to schedule the usage of devices.

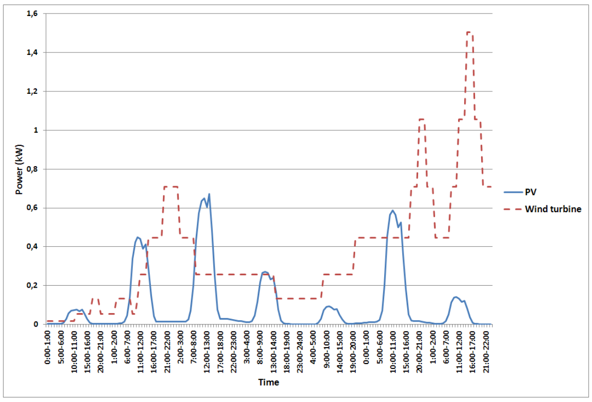


Figure 5.3: Energy forecast for solar power and wind power in the week 5th-11th October 2012.

Naturally, in an implementation of the smart grid services on the field, a reconciliation system to evaluate the gaps between forecast and real prices will be needed a posteriori for billing purposes and to inform the user. The same reconciliation will be required in the actual amount of energy extracted from the grid in case of non exact forecasts of generated renewable energy. This aspect is out of the scope of this work, however we envision that the smart meter will be the center of this reconciliation by storing the in- and out-flow of energy and the tariffs applied at each point in time.

5.3 Smart Grid Aware Buildings

The smart grid simulation tool described in Section 5.2 has been the smart grid engine to realize demand-response functionalities in a real living lab environment at the University of Groningen. A set of software communicating modules have been developed to realize a real demand-response implementation involving real appliances and end users. The components realized are shown in Figure 5.4. In the context of this work we just want to give the essential details and focus on the application of the smart grid functionalities rather than the details of the architecture and the optimization algorithms to schedule the devices. For a more thorough coverage of the architecture and the details concerning the optimization details we refer to [76].

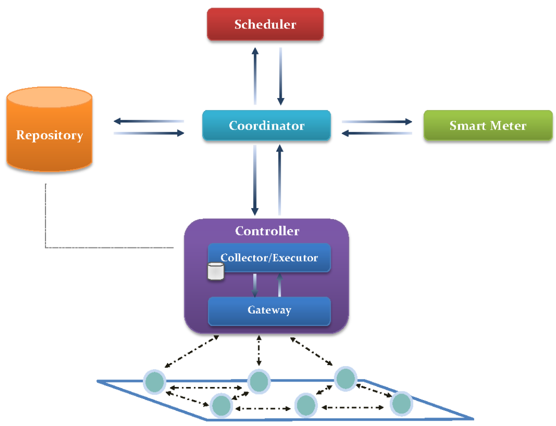


Figure 5.4: smart grid-enabled architecture to implement demand-response fucntionalities.

The smart meter component shown in Figure 5.4 has exactly the functionalities described in Section 5.2. This component is the engine that provides the prices of energy and the amount of power each provider can supply; furthermore, it provides the quantity of power available thanks to the small-scale renewable sources that are supposed to power the office in addition to traditional energy supply. These information, in addition to a set of rules characterizing the operation of the devices (i.e., device policies), are essential to enable the schedule of the appliances with the goal of minimizing energy costs to realize demand-response experiments.

5.3.1 Realization of a smart office

In the facility hosting the Faculty of Mathematics and Computer Science, we have realized a living lab office environment where we have tested the scheduling of a set of appliances (see Figure 5.5). Each of the appliances is connected to the power outlet through a plugwise socket. Plugwise[14] devices are plugs that are attached to a common power outlet on one side and provide a socket to connect an appliance on the other side. A plugwise device is shown in Figure 5.6. The plugwise outlet is able to perform two operations: 1) measure in real-time the amount of power used by an attached device, 2) enable/disable the power flow to the connected appliance acting de facto as an on/off switch. These operations of sense and actuation can be issued to the plugwise sockets via a base station with ZigBee communication capabilities running an appropriate software.

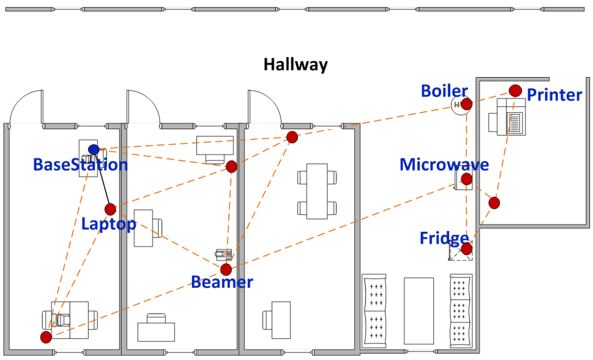


Figure 5.5: Appliances and communication infrastructure in living lab.

For each appliance/device, there is an associated *policy*. A policy is a set of consistent rules that hold for device operations. We have identified these rules by measuring the power consumption of the devices in normal usage conditions and from our usage experience. For example, "a fridge must work at least 15 minutes per hour" to be able to maintain its internal temperature below a certain threshold temperature level. Policies can have different parameters, a few of which are common

[14] `http://http://www.plugwise.com/`

Policy type	Associated device	Description
REPEAT	Fridge, Boiler	Device should be put to a specified state repeatedly with a certain periodicity.
TOTAL	Laptop	Device should operate for at least a certain amount of time.
MULTIPLE	Printer	Device should operate for the time that allows for all scheduled jobs to be performed.
STRICT	Beamer	A strict schedule is given in advance.
PATTERN	Microwave	An expected pattern of device operations.
SLEEP	Any device	No demand for device during the scheduling period.

Table 5.3: Device policies

to all: $(tBegin, tEnd)$ – time period, when the policy is active; and sid – state ID that the policy is applied to. State IDs are unique per device. In general, we assume several possible states per device, together with associated actions to move a device to these states. In the presented setting, each device has two states: "on" and "off," and two associated actions: "turn on" and "turn off."

Figure 5.6: A plugwise Circle plug device.

5.3.2 Benefits of the smart grid-enabled system

In order to have prices for energy that are similar to those available for the end user market, the prices extracted by the price extractor engine of the Smart Grid Simulation Engine have been rescaled (from $/MWh to $/kWh) for this living lab experiment. We have used the system over three weeks in the months of October 2011 and November 2011 performing measurements from Monday to Friday (as in the weekend there is irregular presence). In particular, in the first 2 weeks (W1-W2) we measured energy use in order to define a baseline. The third week (W3), we let the scheduling component control the environment in order to measure the actual savings. Next we present the results in terms of economic savings (due to the varying prices of the smart grid) and of energy savings (due to the introduction of device policies).

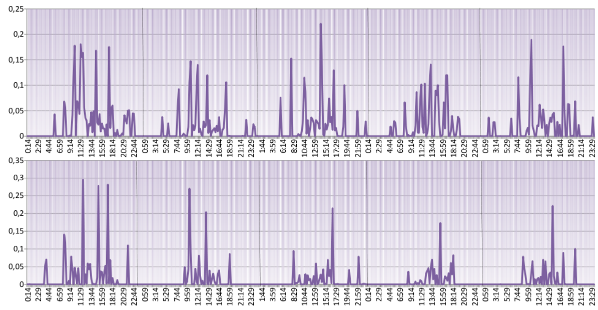


Figure 5.7: Average price ($ per kWh) comparison between non-scheduled (upper chart) and scheduled (lower chart) appliances for each work day.

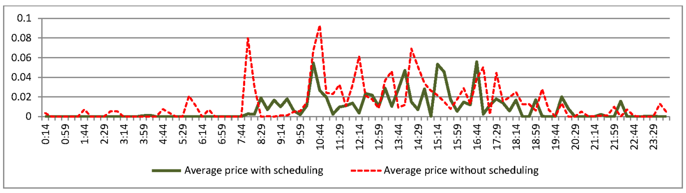
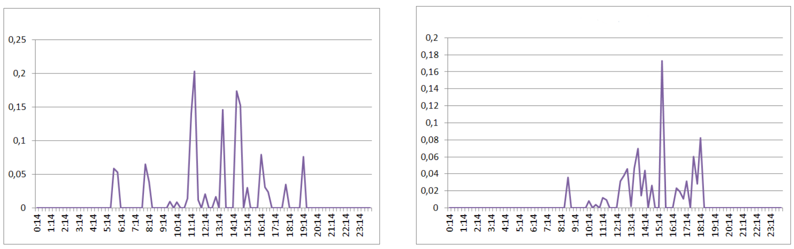


Figure 5.8: Average price ($ per kWh) comparison between scheduled (continuous line) and non-scheduled (dashed line) situation.



(a) Price of energy ($ per kWh) during non-scheduled day October 27^th 2011.

(b) Price of energy ($ per kWh) during scheduled day November 3^rd 2011.

Figure 5.9: Energy price in non-scheduled and scheduled situations.

Economic savings

As we have mentioned before, the goal of the system is to save money for the office taking advantage of the smart grid. Therefore the first evaluation we make is based on taking the energy bill for a week using the developed demand-response system versus a week without it. To make the comparison fair, we use the energy prices of the third week (W3) and apply those same retrieved prices for the energy consumed in the other two weeks of measurement (W1 and W2). The situation between each working day of the two weeks (average) without scheduling policies and the week where the policy has been applied is shown in Figure 5.7, where the cost in term of $ per kWh is shown versus the time of the day (from Monday to Friday). It is interesting to notice the difference in the average price paid for each kWh of energy in the situation without device scheduling and, on the other hand, considering scheduling, the chart is shown in Figure 5.8. On average the price in $ per kWh drops by more than 27% in the two situations. An interesting day where the savings on energy expenses are particularly significant is between the three consecutive Thursdays monitored (October 20^th, 27^th and November 3^rd). Comparing these three days the money savings are on average more than 50%. A comparison between the price paid for energy in each hour between the situation in October 27^th and November 3^rd is shown in Figures 5.9a and 5.9b. In particular, one can see the cut of unnecessary energy expenses related to those consumptions that happen during non-working time (late evening or during the night) by non strictly necessary devices (most notably the hot water boiler). Another optimization the system achieves is the most efficient schedule of devices when the energy generated by photovoltaic panel is more intense and whose cost is generally smaller than energy provisioning on the market.

Energy savings

Although energy saving is not the primary aim of the system, but rather economic savings based on dynamic pricing, the use of policies for devices alone provides for energy saving in absolute terms. Figure 5.10 shows the average energy consumption (kWh) considering the use and the absence of the scheduling system. One can see that the scheduling reduces the consumption of devices that are not used during non working hours and that do not impact the habits of the user (e.g., keeping hot water boiler working at night); in addition the scheduler tries to use at best the cheap electricity coming from the solar panels during day-light hours. Figure 5.11 visually reinforces the idea of reducing loads when unnecessary among the normal (upper chart) and the scheduled solution (lower chart): one notices a more compact chart in which energy is mostly used during daytime (8 a.m. 6.30 p.m.) in each day of the week. On average the savings in energy consumed between the situation without the scheduling policy and the situation considering it, is more than 15%.

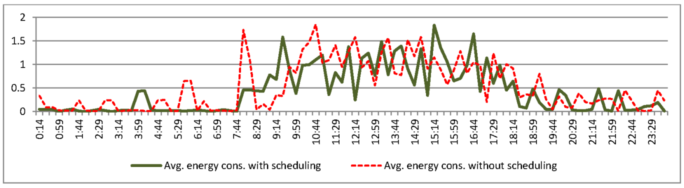


Figure 5.10: Average energy usage (kWh) comparison between scheduled (continuous line) and non-scheduled (dashed line) situations.

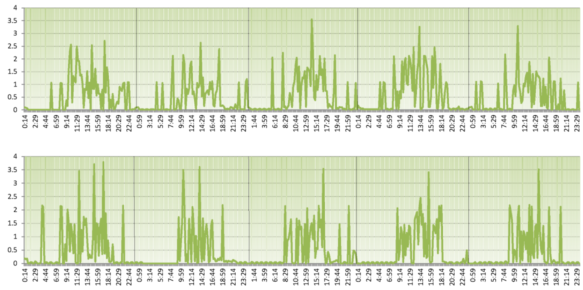


Figure 5.11: Energy (per kWh) comparison between non-scheduled (upper chart) and scheduled (lower chart) appliances for each work day.

5.4 Agent-Based Energy Market

Our vision for the future is a completely free energy market. In this market consumers, prosumers and traditional energy companies can participate offering energy directly on a short term time horizon (e.g., 15 minutes to 1 hour). In addition, we envision that the buildings of the future will be aware of their energy consumption and of the energy tariffs through their energy management systems. Buildings will regulate their energy use and appliance operations based on the price signals for energy coming from the smart grid as we showed in Section 5.3. However, in a totally free market the building (via its energy management system) has to provision for its energy consumption and access directly to the energy market. Other buildings and actors of the smart grid with generating capabilities will do the same creating a market where energy is traded. In order to test this concept and interactions we have realized a software solution based on agent technology to test the feasibility of this kind of energy market on a real software platform able to interact with real smart meter devices too. Smart meters are likely to be an essential component of the future energy management system of buildings. Our second aim is to test the existence of procurement strategies that minimize the average price paid for energy by consumers.

5.4.1 Roles in the smart grid energy market

In the market, we distinguish a number of roles. The fundamental actors are: few traditional big generating companies (i.e., gencos), a limited number of prosumers that are able to produce small quantities of energy (compared to gencos) and a high number (compared to gencos and prosumers) of buyers that are interested in purchasing energy at the cheapest price possible. The market also contains a third party authority, known as balancer. This authority has a role similar to the role today played by ISOs or RTOs. In fact, the balancer has information about the quantities of energy produced and demanded by the various parties. Its role is to act as a intermediary with the aim of maintaining as much as possible the energy balance equilibrium on the power grid.

In such a model, the described actors have, possibly conflicting, goals. The prosumers have the goal to sell any surplus power on the energy market at the highest possible price. We consider that the energy produced by prosumers has a price usually significantly lower than the price offered by the gencos. In addition, the main issue is that the power provided by prosumers is not sufficient to supply the whole demand, but only a fraction of it. We consider this condition to be realistic given the adoption rate of small-scale distributed energy resources. The gencos have the goal to sell energy optimizing price per unit, that is, since production costs do not grow linearly, they want to sell energy at the price yielding the highest possible revenue for them [142]. The balancer has the goal to keep (a geographically identified portion of) the market in balance to avoid failures on the grid. Therefore, the amount of demand not met by the offer by the prosumers should be backed up by the gencos. The external balancer entity calculates the sum between every capacity of every prosumer and forecasts the amount of extra energy that needs to be produced by the gencos to satisfy the expected demand. Weather conditions are essential information for the balancer. To compute the power equilibrium condition the balancing authority uses an equation showing the relation between the total energy demand of an area and the supply capacity of the sellers for that area. There are several mathematical models to describe balancing equations (e.g., [109, 112, 202]), though all have the same basic underlying idea which is to set to zero the algebraical sum between energy demand and supply.

Let $D_{\Delta t}$ be the sum of all the demands of all consumers in a given time interval, we define its balancing equation as

$$D_{\Delta t} = \sum_{k=1}^{C_{\Delta t}} D_k = \sum_{i=1}^{N_g} S_{\Delta t}^{Gc_i} + \sum_{j=1}^{M_p} S_{\Delta t}^{Pr_j} \tag{5.4}$$

where:

- $C_{\Delta t}$ is the number of consumers at time interval Δt;
- D_k is the energy demand by the k^{th} energy buyer;
- $S^z_{\Delta t}$ is the energy supply provided by z^{th} source of energy at time interval Δt;
- N_g is the total number of number of gencos;
- Gc_i is the i^{th} genco;
- M_p total number of prosumers; and
- Pr_j is the j^{th} prosumer.

The equation simply states that the sum between the two different production sources (gencos and prosumers) should be equal to the total consumer demand. This is a simple equation that does not take into account other aspects such as energy losses in transportation and production, forecasting errors in demand, nor any quality of services of the energy producers. However, for the purpose of this work more focused on realizing the platform and the market, the relaxation of all the details of the real situation are adequate.

5.4.2 Agent modeling

Agents are software programs that are programmed to act with a specific purpose on behalf of an authority or user or role [91]. Therefore, software agent technology is the perfect match to implement the smart grid energy market given the several actors and roles given the description above. We distinguish between main agents and auxiliary agents. The former represent energy consumers and energy generators, whereas the latter do not directly deal with the energy purchase and sell, but provide information and mediation to support the behavior of the main agents.

In synthesis the main agents are:

- *Consumers*: their goal is to buy energy.
- *Prosumers*: their goal is to produce and consume electricity being able to both buy and sell it. They can produce a limited quantity of energy (compared to a genco) thanks to small-scale energy production devices such as small wind turbines, solar panels, and micro-combined heat and power (micro-CHP).
- *Gencos*: their goal is to sell energy at various scale. They are the traditional big energy generating companies which can be (or not) also in charge of the distribution of the energy.

As for the auxiliary agents, we initially identify three:

- *Balancer*: it has a mediating role among main agents to keep the balance of supply and demand.
- *Time Agent*: it defines the starting and closing of a time interval for negotiation (i.e., Δt interval in Equation 5.4).
- *Weather Agent*: it provides localized information about weather forecast, especially parameters useful to predict small-scale generation.

In addition to these agents, for the simulation purposes two more agent roles need to be defined since we resort to the Java Agent Development Environment (JADE) [44]. The first agent created is the Creator Agent that generates all the other agents. Another essential agent that is present in the architecture is the Directory Facilitator Agent (DF) that provides a yellow pages-like service for the platform.

5.4.3 Agent interactions

All the interactions between agents take place by message exchanges. The message exchanges need to be standardized in order to make sure that any new agent can join an existing infrastructure. We chose to go for the most accepted standard for agent communication, the Foundation for Intelligent Physical Agents (FIPA) standard [155]. We consider two kind of messages: synchronization messages and market related messages.

From the implementation point of view, the chosen platform (i.e., JADE) follows an asynchronous message passing paradigm and it is based on the Agent Communication Language (ACL), while being fully FIPA compliant. Every agent has a queue that stores the incoming messages, the extraction of the message from the queue is then up to the programmer.

The essential parts composing a message according to the FIPA/ACL language are:

- *Content* is the payload of the message itself.
- *Performative* represents the purpose of the message, or better the communicative intention. Independently from the content, an agent could decide to try to better understand what a sender is saying by just fetching this part of the message. In particular, for the smart grid energy market, we consider the following constructs:

  - INFORM and INFORM_REF performatives represent generic informative messages used for synchronization purposes. The latter is used as synchronization signal for final operations (e.g., a buyer informs a balancer that he has stipulated a contract successfully).
  - PROPOSE, REFUSE and ACCEPT_PROPOSAL performatives are used in market related messages. The first one is used for a bid proposal sent by a buyer agent or it provides information about the starting price of a negotiation when sent by a energy seller agent. The second performative is used to refuse a single offer, while the last performative is used to accept an offer proposal.
  - CANCEL performative is used by a prosumer in order to abort the current auction with the specified buyer.

- *Language* represents the syntax used to express the content, however this aspect is not used in our project.

- *Ontology* represents the knowledge representation of the domain. This aspect is not used in the current version of our project, though might be considered in future extensions.

5.4.4 Agent behavior in the market

In the domestic energy market, all the agents have a specific goal that is, trade a good and maximize profit. Therefore, the agents are competing against each other for obtaining the energy at the best price and selling it at the highest one. The actual trade is based on an auction system in which all main agents participate at every round of trading. Thus the situation is as follows.

- A buyer can stipulate a contract with a prosumer after winning an auction round, realized with sealed bids; every buyer can send several offers to suitable prosumers in any given round.

- Prosumers' energy starting price is considerably lower than gencos' initial contract prices. The underlying assumption is that any positive amount is helpful in the return from a sunk investment on renewable energy generation.

- Weather conditions during an observed interval can prevent a prosumer to generate enough electricity to be sold.

- Prosumers communicate to buyers an initial starting price that is influenced by contracts with energy transmission and distribution operators and a random cost due to the devices used to produce electricity (e.g., maintenance costs).

- The energy produced by a prosumer has to be sold quickly and cannot be stored or buffered for selling at a later auction round.

Every prosumer is in direct competition with other energy sellers therefore they propose an appealing starting price and make an intelligent use of refusing bids in order to rise the price without letting buyers contact other sellers. On the other hand, gencos are big energy generating companies and they have a theoretical infinite amount of energy supply, but they also have some peculiarities to be taken into account: i) gencos sell energy with contracts lasting for one time interval or more at a given fixed price; ii) gencos contracts can be stipulated much faster since there is no auction process; iii) gencos prices are in general higher than prosumers' starting prices; iv) gencos prices depend on exceeding production threshold known a priori. This last aspect implies that energy exceeding the threshold will be more expensive for the genco and thus for the buyer. This additional cost models the operation of extra expensive power plants to compensate an excessive demand. Therefore, the unit price of energy the gencos sell on the market can be represented by the following threshold function:

$$E_{uc} = \begin{cases} Cost_{energy} \text{ if below supply threshold} \\ Cost_{energy} \times (EC \times A) \text{ otherwise} \end{cases} \tag{5.5}$$

where $EC > 1$ is an external costs constant and $A \in \mathbb{N}_1$ is the number of energy units above the threshold. In other words, asking gencos energy contracts when the threshold is already exceeded leads to more expensive contract prices. Those prices rises as we get further from the specified threshold.

In these conditions, the challenge is strictly related with the ability of agents to quickly obtain an energy contract either with a prosumer or with a genco since as time passes it is more likely that gencos will exceed their production threshold thus increasing the price of a contract. In fact, if every buyer follows a natural strategy that suggest him to contact the prosumer first hoping in cheap contracts. The buyer is then involved in spending time in the auction process with offers and bids rising. If no contract with a prosumer is found it is likely that the gencos in the meanwhile have exceeded their threshold and only high priced energy is available.

A modeling of the best strategies for the agents, is beyond the scope of the present treatment which aims at providing a modeling and a software platform for the domestic market. We remark, however, that the family of minority games is a good way of representing the interaction and finding suitable agent strategies in such a market. In minority games, two different behaviors for a single player are admitted and s/he wins if s/he chooses the path taken by the minority of the players. Applying this game to the domestic energy market means that each buyer agent can learn and adjust its behavior (i.e., contacting a prosumer or a genco) reducing the average cost of energy paid by the buyers [35]. In particular, the most similar game theoretic approach that presents a situation similar to the one just explained for the energy market model developed is the *El-Farol-Bar* problem [69]. It is possible to use the solution identified for the *El-Farol-Bar* game by Whitehead [224] and extend it to take into account the even higher level of complexity of the energy market model presented here [35].

5.5 Bits and Electrons Hand in Hand

In this chapter, we have seen how ICT is essential for the realization and success of the smart grid. To add value for the end user and make energy provisioning and distribution more efficient, effective, and sustainable ICT must be at the core of the smart grid to enable new scenarios of energy use. ICT opens new services and opportunities in the energy world. First of all, by using flexible information oriented paradigms such as SOA, it is easier and more efficient to share information between the many actors in the smart grid panorama. SOA is also an approach able to easily and fast adapt to the changes that a still uncertain framework such as the smart grid requires. Standards are under development and test of solutions are underway. SOA has also the capability to adapt solutions easily especially in heterogeneous environments such as the power industry of today and tomorrow where more technologies have to be integrated in a final solution. We have shown through this chapter how today some of the ingredients to realize the demand-response service for the smart grid are available. By just providing little adaptations, the ingredients can be used in real smart grid service contexts. In our tests of an office connected to the smart grid, we have shown that real benefits for the end user are possible in monetary terms, and at the same time more renewable sources can be used as energy production sources. ICT opens also new scenarios and possibilities such as prosumer based energy markets. Users can interact by using software applications that work on their interest (i.e., agents based) and automatically trade energy to achieve profits (in the prosumer case) or lower the energy bill (in the consumer case). Proof of concepts such as the one realized in our faculty, or small living labs (e.g., PowerMatching City Hoogkerk [22]) show the feasibility of smart grid ideas and scenarios. The efforts are now devoted to scale these proofs of concept to country-wide solutions and investigate additional services. These new ICT services have the aim of improving energy efficiency and increment the penetration of renewable sources.

Chapter 6

Conclusion

We envision a more sustainable, reliable, efficient, decentralized, and affordable energy supply where users can provide for their own supply and exchange energy in a free energy market at local scale, such as neighborhood or city level. This vision can be reconciled with the various meanings [139] and visions [208] of the smart grid. Fostered by this inspiration, we have studied a new approach to the electrical grid, both in its infrastructural characteristics proposing evolutions, and in realizing applications for the future energy services and market.

6.1 The Emergence of a New Grid

The smart grid is not yet implemented nor standardized and still misses a unique definition. In essence, it deals with enhancing the electrical system with the functionalities of Information and Communication Technologies (ICT) to achieve more flexibility, more resilience, and more efficiency. The aim is to implement a high share of renewables in the energy generation sources and provide the end user with even a better service (e.g., more reliable, better tariffs) than the traditional grid.

Our work focuses on two aspects: (i) the modeling and evolution of the distribution grid to better accommodate the smart grid vision using the complex network analysis approach and (ii) software and applications in the smart grid context.

Considering the first topic of research, we have focused on the medium and low voltage layers of the power grid infrastructure, that is, the part of the grid that is closer to the end user. The novelty of the work has been severalfold. A first novelty lies in looking at the medium and low voltage grid that is a layer of the electrical systems that has not received much attention. A second novelty consists of analyzing the power grid using complex network analysis taking into account the physical parameters of topology that influence the price of electricity through a new set of topological metrics. A third one consists of proposing novel smart grid related metrics to evaluate the topology of the grids. A fourth point deals with using complex network as a design tool for the design and expansion of the grid by defining a set of metrics for the smart grid. Fifth, this approach sets the base for a decision support system to guide the planning and evolution of the grid towards a smart grid where local energy exchange is prominent. Topological studies of the power grid have so far focused only on the high voltage grid and almost always in a traditional conception (i.e., not related to the smart grid) of the power grid in order to understand failures related to topological properties. Our work has focused on the medium and low voltage network and we have investigated the topological properties of physical samples belonging to the Dutch grid. In particular, we have been able to associate topological measures with the price of distributing electricity and compare the different Dutch topologies according to those measures. In the literature, complex network approaches have been mainly used to analyze networks in order to grasp the fundamental properties of the system represented in the network. These fundamental properties are extracted through measures and indicators of the whole network. Another effort in the complex network community has been to create models to mimic the dynamics of network evolution systems (e.g., the Barabási-Albert preferential attachment model). We have used complex network tools also to design medium and low voltage grids for the future energy system by considering a set of properties the networks have to satisfy. Our investigation focused on both models coming from complex network analysis literature which might be used for planning the infrastructures in settlements to be realized (e.g., developing countries), and on the evolution of current distribution grid topologies as we do for the Dutch samples. Considering the whole work on the grid infrastructure side, we have set the basics of realizing a decision support system for analysis and planning of the smart grid. A skeleton diagram of the envisioned decision support system is shown in Figure 6.1. In the figure, several phases and input are considered to plan the evolution of the infrastructure where a local energy exchange is the overarching goal. It starts with a pre-processing phase where the input data of the grid is converted into a graph; the output of this initial phase is a Power Grid graph. The following phase consists of the analysis of the topological properties characterizing the graph. The output of this phase consists of a set of values representing the metrics related to the power grid that influence the price of electricity (α and β metrics in the figure). The process continues with the generation of a network model. The number of nodes and edges of this reference model are provided according to the targets for the cost-related parameters (α and β) and the will to invest by the stakeholders. Based on the theoretical model identified, the physical network under assessment is then fitted to a topological structure similar to the one of the model. Several solutions are provided that differ in the topology and the α and β metrics. All of these solutions are then input into a computer-assisted decision support system that presents the benefits/costs of the evolution of the network. An expert is involved in the selection of the most promising candidates for the evolution process that is built by the computer because there are other political, geographical, and legal considerations that may rule in favor or against any possible evaluation. Once the decision is made, the adaptation of the physical grid can begin. We have realized several steps of this system, but it is still not a fully automated piece of software.

Considering the second topic, we have looked into the software approach for an heterogeneous environment such as the smart grid and built scenarios and applications for the future. Our finding, in line with other recent proposals, e.g. [221, 53], is that software-oriented architectures (SOAs) are an ideal candidate for integrating different systems coming from traditional power engineering telecontrol (e.g., supervisory control and data acquisition) and new systems featuring smart grid functionalities (e.g., dynamic energy pricing, demand-response). The applications that we have realized go into the direction of exploring scenarios and services for the smart grid. On one side, we have tested appliance scheduling and automation in a dynamic pricing environment with realistic energy prices and renewable energy production in an office environment with real users. On the other side, another application considered an essential ingredient of the smart grid in our vision: a free market where consumers, prosumers and power companies can participate. We have tested agent technologies to enable such a scenario coupled with real smart meter devices and identified a promising strategy for the users to supply their electricity based on minority games.

6.2 Open Issues and Future Directions

The smart grid is an evolving subject with test beds and pilot projects appearing all over the world. The subject investigated in this thesis is open to further research and improvements both on the grid side and on the software/application side.

Considering the power grid, several aspects can be considered to follow up. First of all, in order to move from a theoretical high level planning and decision support system to a more fine grained planning tool close to traditional power engineering, power flow analysis should be interpreted. With more digital reading, precise information on the power flows in the medium and low voltage grid will be available. The goal is to evaluate the performance of the synthetic networks proposed and the Dutch grid topologies evolved using complex network analysis principles. Such analysis could give a better insight into the real application on the field of our findings. Along the same lines, we propose the evaluation of the findings of this thesis with engineers from the power industries to understand how to improve the decision support systems skeleton that we have developed so far into a fully usable tool to test and plan distribution grid evolution for the smart grid.

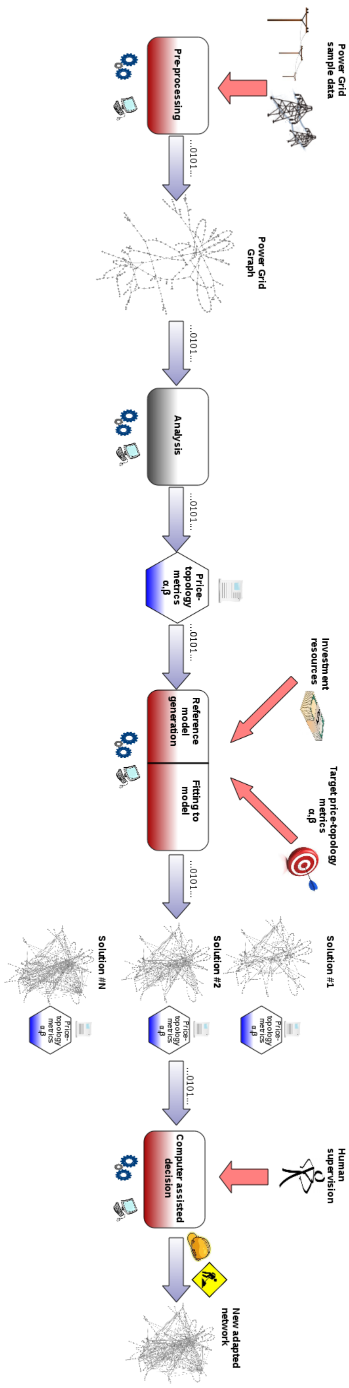


Figure 6.1: Decision support system for smart grid planning.

Other research topics involve the identification or design of topological models tailored to the smart grid. The models used for the synthetic generation (in Chapter 4) were based on existing complex network analysis. For the smart grid, a specific model could be developed that enables topological and spatial efficiency and robustness on the one hand, and that is electrically valid on the other hand; small-world networks seem as a good starting point for this further investigation. The decision support system depicted in Figure 6.1 needs refinement. Especially the translation from economic targets to topological targets has to be fine-tuned to make a translation in quantitative objectives. An open issue is the influence on the electricity price paid by the end user versus the economic advantage of end user topologies. A detailed economic analysis should be performed to understand the impact of topologies that should reduce the price of electricity distribution and the effects on the energy bills of the end users. Last, but not less important, is the comparison of the analysis and evolution proposed for the Dutch grid to other distribution grid topologies coming from other geographies and countries.

Considering the application and ICT side of the smart grid, two lines of improvement of the software realized are foreseeable. On the engineering side, the developed scheduling system should be reduced to its essential components and embedded in a home energy management system as well as the agent technology application for market interaction. This is more of an industrial research and developmental issue. The software system should also be enhanced with artificial intelligence and machine learning techniques to learn user preferences and be aware of external conditions in order to act or react. A series of tests on home energy management system and intelligent appliances in the home environment should be performed in collaboration with sociologists and psychologists to obtain the maximum results and feedback on the approach to energy savings and products when dealing with end users. The future of services involving the smart grid have to be broad (i.e., large number of services available) and personalized for the user (e.g., dedicated tariffs based on user profile and his building). We envision cloud computing services that provide the user with personalized information where the back-end of these services is based on the SOA approach. Tests and scaling of these applications applied in the power domain are another field of research and development.

There are other open points that go beyond the technical and ICT aspects of the smart grid that, anyway, must be considered especially on the influences that they might pose on the evolution of the smart grid. From a sociological side, open points are the acceptance of the smart grid by the end users accustomed to the traditional way of dealing with the electrical system. Another aspect from the psychological side that deserve attention concerns the incentives to provide to the users to implement energy efficiency policies or demand-response initiatives. Price alone could not be the best to approach all the segments of the population [23]. Open points exist in other fields, such as law. A full unbundling such as the one we envision, have to pass through the legal institution that have to modify the current legal system and maybe create new subjects with legal responsibilities over the operations of the local grid (today this is in place for the high voltage grid). Economists and policy makers are also investigating the pricing mechanisms to be applied in a scenario with more distributed generation and potentially more small markets that are economically viable and at the same time that do not have to impact on the stability of the electrical system.

We have found that complex networks methodology prove to be a good tool to analyze and design networks. However, given the peculiarity of the physical properties that lay in the electrical system, a more detailed model of the lines need to be considered. In particular, the analysis of power flows is important to translate the theoretical models to networks that can be later engineered. Other interesting ways of evolving networks are, for instance, through multi-objective optimization on complex networks or complex network evolution using a genetic algorithm approach. We also consider the scientific discipline of complex systems as an interesting framework to use in the study of smart grid, given the studies performed and the results obtained in the study of city and infrastructure scaling [21]. The complexity of the smart grid lies not only in the technical aspects of a system, but also in the social dynamics overlying the system: user interactions and their attitude to buy or sell energy on the grid, the mutual influence through social media modifying their energy behavior might have effects on the electrical system (macro) not predictable by the analysis of the single individual (micro) [27]. On the software side, we have seen that the drivers are available to realize demand-response and appliance automation functionalities related to the smart grid. The aspect that deserves more attention is the involvement of the users, their feedbacks and their reactions to the new electrical system. Implementation and tests on large-scale are the necessary steps to proceed towards smart grid realization.

Obviously, all of the speculations for future direction of research and development associated with the topics described in this chapter depend heavily on the transition towards the smart grid. The path is, in the broad sense, set. Different countries are adopting different approaches at different speeds; the adoption is influenced by the natural resources available, the energy policies, and the economy involved in the energy sector. Until new solutions for harvesting clean energy become available, the smart grid is the solution on the horizon that can enable the transition towards a sustainable future providing high integration of distributed renewable energy sources into the system with higher flexibility and enabling new opportunities for energy utilities.

Appendix A

Graph Theory and Complex Network Fundamentals

A.1 Graph Theory and Complex Network Definitions and Properties

The approach used in this thesis to model the power grid and its evolution is based on Graph Theory and Complex Networks Theory. Here we recall the basic definitions that we use throughout the thesis and refer to standard Graph Theory textbooks such as [24, 25] or those focusing on Complex Networks such as [153, 49] for a deeper introduction.

First, for the sake of completeness we define once again a graph for the power grid and its weighted representation as we described in Chapter 3.

A.1. DEFINITION (GRAPH). *A* graph *G is a pair of sets G(V,E) where V is the set of vertexes and E is the set edges. An edge $e_{i,j}$ is a pair of vertexes (v_i, v_j). If $(v_i, v_j) \in E$ then v_i and v_j are said to be* adjacent *or* neighboring *and are called* end-vertexes *of the edge.*

A.2. DEFINITION (POWER GRID GRAPH). *A* Power Grid graph *is a graph $G(V, E)$ such that each element $v_i \in V$ is either a substation, transformer, or consuming unit of a physical power grid. There is an edge $e_{i,j} = (v_i, v_j) \in E$ between two nodes if there is physical cable connecting directly the elements represented by v_i and v_j.*

One can also associate weights to the edges representing physical cable properties (e.g., resistance, voltage, supported current flow).

A.3. DEFINITION (WEIGHTED GRAPH). *A* weighted graph *is a pair $G(V, E)$ where V is the set of vertexes and E is the set of edges. An edge $e_{i,j,w} = (v_i, v_j, w)$ is a triple where $v_i, v_j \in V$ and $w \in \mathbb{R}$. w is called* weight *of the edge.*

A.4. DEFINITION (WEIGHTED POWER GRID GRAPH). *A* Weighted power grid graph *is a Power Grid graph $G_w(V, E)$ with an additional function $f : E \to \mathbb{R}$ associating a real* *number to an edge representing the physical property of the corresponding cable (e.g., the resistance, expressed in Ohm, of the physical cable).*

A first classification of graphs is expressed in terms of their size.

A.5. DEFINITION (ORDER AND SIZE OF A GRAPH). *Given the graph G the order is given by* $N = |V|$*, while the size is given by* $M = |E|$.

From *order* and *size*, it is possible to have a global value for the connectivity of the vertexes of the graph, known as *average node degree* . That is $\langle k \rangle = \frac{2M}{N}$. To characterize the relationship between a node and the others it is connected to, the following properties provide an indication of the bond between them.

A.6. DEFINITION (ADJACENCY, NEIGHBORHOOD AND DEGREE). *If* $e_{x,y} \in E$ *is an edge in graph G, then* x *and* y *are* adjacent, *or* neighboring, *vertexes, and the vertexes* x *and* y *are* incident *with the edge* $e_{x,y}$. *The set of vertexes adjacent to a vertex* $x \in V$*, called the* neighborhood *of* x*, is denoted by* Γ_x. *The number* $d(x) = |\Gamma_x|$ *is the* degree *of* x.

When considering a weighted graph it is possible to extend the definition of degree to have an idea of the importance of a node based on the incident edges.

A.7. DEFINITION (WEIGHTED DEGREE). *Let* $x \in V$ *be a vertex in a weighted graph G, the* weighted degree *of* x, $d_w(x)$ *is:*

$$d_w(x) = \sum_{y \in \Gamma(x)} w_{x,y}$$

where $w_{x,y}$ *is the weight of the edge joining vertexes* x *and* y*, and* $\Gamma(x)$ *is the neighborhood of* x.

A measure of the average 'density' of the graph is given by the clustering coefficient, characterizing the extent to which vertexes adjacent to any vertex v are adjacent to each other.

A.8. DEFINITION (CLUSTERING COEFFICIENT (CC)). *The* clustering coefficient γ_v *of* Γ_v *is*

$$\gamma_v = \frac{|E(\Gamma_v)|}{\binom{k_v}{2}}$$

where $|E(\Gamma_v)|$ *is the number of edges in the neighborhood of* v *and* $\binom{k_v}{2}$ *is the total number of* $possible$ *edges in* Γ_v.

This local property of a node can be extended to an entire graph by averaging over all nodes.

Another important property is how much any two nodes are far apart from each other, in particular the minimal distance between them or shortest path. The concepts of *path* and *path length* are crucial to understand the way two vertexes are connected.

A.9. DEFINITION (PATH AND PATH LENGTH). *A* path *of G is a subgraph P of the form:*

$$V(P) = \{x_0, x_1, \ldots, x_l\}, \qquad E(P) = \{(x_0, x_1), (x_1, x_2), \ldots, (x_{l-1}, x_l)\}.$$

such that $V(P) \subseteq V$ *and* $E(P) \subseteq E$. *The vertexes* x_0 *and* x_l *are* end-vertexes *of P and* $l = |E(P)|$ *is the* length *of P. A graph is* connected *if for any two distinct vertexes* $v_i, v_j \in V$ *there is a finite path from* v_i *to* v_j.

A.10. DEFINITION (DISTANCE). *Given a graph G and vertexes* v_i *and* v_j*, their* distance $d(v_i, v_j)$ *is the minimal length of any* $v_i - v_j$ *path in the graph. If there is no* $v_i - v_j$ *path then it is conventionally set to* $d(v_i, v_j) = \infty$.

A.11. DEFINITION (SHORTEST PATH). *Given a graph G and vertexes* v_i *and* v_j *the* shortest path *is the the path corresponding to the minimum of to the set* $\{|P_1|, |P_2|, \ldots, |P_k|\}$ *containing the lengths of all paths for which* v_i *and* v_j *are the end-vertexes.*

A global measure for a graph is given by its average distance among any two nodes.

A.12. DEFINITION (AVERAGE PATH LENGTH (APL)). *Let* $v_i \in V$ *be a vertex in graph G. The* average path length *for G* L_{av} *is:*

$$L_{av} = \frac{1}{N \cdot (N-1)} \sum_{i \neq j} d(v_i, v_j)$$

where $d(v_i, v_j)$ *is the finite distance between* v_i *and* v_j *and N is the order of G.*

A.13. DEFINITION (CHARACTERISTIC PATH LENGTH (CPL)). *Let* $v_i \in V$ *be a vertex in graph G, the* characteristic path length *for G,* L_{cp} *is defined as the median of* d_{v_i} *where:*

$$d_{v_i} = \frac{1}{(N-1)} \sum_{i \neq j} d(v_i, v_j)$$

is the mean of the distances connecting v_i *to any other vertex* v_j *in G and N is the order of G.*

When considering a weighted graph the definition of CPL can be easily extended to account for weights characterizing the edges.

A.14. DEFINITION (WEIGHTED CHARACTERISTIC PATH LENGTH (WCPL)). *The* weighted characteristic path length *for graph G, L_{wcp} is the median for all $(v_i, v_j) \in V$ of d_{w_i} where:*

$$d_{w_i} = \frac{1}{(N-1)} \sum_{i \neq j} e_{w_{i,j}}$$

is the mean of the weighted distances connecting v_i to any other vertex v_j and $e_{w_{i,j}}$ is an edge in the minimal weighted path between v_i and v_j in G and N is the order of G.

To describe the importance of a node with respect to minimal paths in the graph, the concept of betweenness helps. Betweenness (sometimes also referred as *load*) for a given vertex is the number of shortest paths between any other nodes that traverse it.

A.15. DEFINITION (BETWEENNESS). *The* betweenness $b(v)$ *of vertex $v \in V$ is*

$$b(v) = \sum_{v \neq s \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where $\sigma_{st}(v)$ is 1 if the shortest path between vertex s and vertex t goes through vertex v, 0 otherwise and σ_{st} is the number of shortest paths between vertex s and vertex t.

Looking at large graphs, one is usually interested in global statistical measures rather than the properties of a specific node. A typical example is the node degree, where one measures the node degree probability distribution.

A.16. DEFINITION (NODE DEGREE DISTRIBUTION). *Consider the degree k of a node in a graph as a random variable. The function*

$$N_k = \{v \in G : \ d(v) = k\}$$

is called probability node degree distribution*.*

The shape of the distribution is a salient characteristic of the network. For the power grid, the shape is typically either exponential or a power-law [15, 6, 159, 185]. More precisely, an exponential node degree (k) distribution has a fast decay in the probability of having nodes with relative high node degree. The relation:

$$P(k) = \alpha e^{\beta k}$$

follows, where α and β are parameters of the specific network considered. On the contrary, a power-law distribution has a slower decay with higher probability of having nodes with high node degree. It is expressed by the relation:

$$P(k) = \alpha k^{-\gamma}$$

where α and γ are parameters of the specific network considered. We remark that the graphs considered in the power grid domain are usually large, although finite, in terms of *order* and *size* thus providing limited and finite probability distributions.

A Graph can also be represented as a matrix, typically an adjacency matrix.

A.17. DEFINITION (ADJACENCY MATRIX). *The adjacency matrix $A = A(G) = (a_{i,j})$ of a graph G of order N is the $N \times N$ matrix given by*

$$a_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

A graph can be also represented by its Laplacian matrix:

A.18. DEFINITION (LAPLACIAN MATRIX). *Let $D = (D_{ij})$ be a diagonal matrix with $D_{ii} = d(v_i)$ the degree of vertex v_i in graph G and A the adjacency matrix of G. The matrix $L = D - A$ is the Laplacian matrix of graph G.*

A.2 Graph properties example

For the sake of clarity, we apply the definitions introduced in Section A.1 to a simple graph example and compute the main properties that characterize it. Figure A.1 shows the simple graph that we consider.

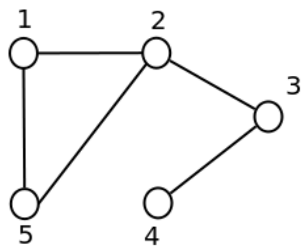


Figure A.1: A simple graph.

Order* and *Size The graph G shown in Figure A.1 is characterized by the set of vertexes V:

$$V = \{1, 2, 3, 4, 5\}$$

and by the set of edges E:

$$E = \{(1, 2), (1, 5), (2, 3), (2, 5), (3, 4)\}$$

the *order* and *size* of G are $|V| = 5$ and $|E| = 5$ respectively.

Neighborhood The neighborhood of vertex 1 in Figure A.1 is:

$$\Gamma(1) = \{2, 5\}$$

its degree is:

$$d(1) = |\Gamma(1)| = 2$$

The average node degree for G is:

$$\langle k \rangle = \frac{2 \cdot 5}{5} = 2$$

Clustering The clustering coefficient for vertex 1 is:

$$\gamma_1 = 1$$

while for the entire graph it is the average of $\{\gamma_1 = 1, \gamma_2 = \frac{1}{3}, \gamma_3 = 0, \gamma_4 = 0, \gamma_5 = 1\}$ that is:

$$\gamma_G = 0.467$$

Path, Distance and Path Length Example of paths between vertex 1 and vertex 4 are:

$$P_{1,4} = \{1 - 2 - 3 - 4\}$$

whose length $l_P = 3$ but also:

$$P^{'}_{1,4} = \{1 - 5 - 2 - 3 - 4\}$$

is a valid path whose length $l_{P'} = 4$, therefore the shortest path between vertex 1 and vertex 4 is $P_{1,4}$.

The distance between vertex 1 and vertex 4 is:

$$d(1, 4) = 3$$

while distance between vertex 1 and vertex 5 is:

$$d(1, 5) = 1$$

Average and Characteristic Path Length The average path length is given by:

$$L_{AV} = \frac{1}{5 \cdot 4}(7+5+6+9+7) = \frac{34}{20} = 1.7$$

The characteristic path length is the median of:

$$\{d_{v_1} = \frac{7}{4}, d_{v_2} = \frac{5}{4}, d_{v_3} = \frac{6}{4}, d_{v_4} = \frac{9}{4}, d_{v_5} = \frac{7}{4}\}$$

that is:

$$L_{CP} = \frac{7}{4} = 1.75$$

Betweenness Vertex 2 is involved in the following shortest paths:

$$P_{1,3}, P_{1,4}, P_{3,1}, P_{3,5}, P_{4,1}, P_{4,5}, P_{5,3}, P_{5,4}$$

therefore betweenness of vertex 2 is:

$$b(2) = 8$$

Matrix Representation of a Graph The Adjacency matrix for graph G is:

$$A(G) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

The Laplacian matrix for graph G is:

$$L(G) = \begin{pmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ -1 & -1 & 0 & 0 & 2 \end{pmatrix}$$

Appendix B

Complex Network Models

Network models and the algorithms to build them are essential in knowing how to generate new networks or how to re-arrange connections in already available networks to obtain networks that satisfy a specific set of properties. Here we give a brief description of the models used in this work that is based on the more complete and thorough description in [39] and [151] to which we refer for more details.

B.1 Building Synthetic Networks

Random Graph

A Random Graph is built by connecting each pair of nodes with an edge with probability p. It is due to the pioneering studies of Erdős and Rényi [67]. More precisely, there are two ways to built a Random Graph, (a) the $G_{N,p}$ model proposed by Erdős and Rényi considers a set of N nodes and for each pair of nodes an edge is added with a certain probability p; (b) the $G_{N,M}$ model considers with equal probability all the graphs having N vertexes and exactly M edges randomly selected among all the possible pairs of edges. The models have the same asymptotic properties. In this work we use the $G_{N,M}$ model since we are interested in setting both the number of nodes and edges for the networks to generate. A Random Graph with *order* 199 and *size* 400 is shown in Figure B.1.

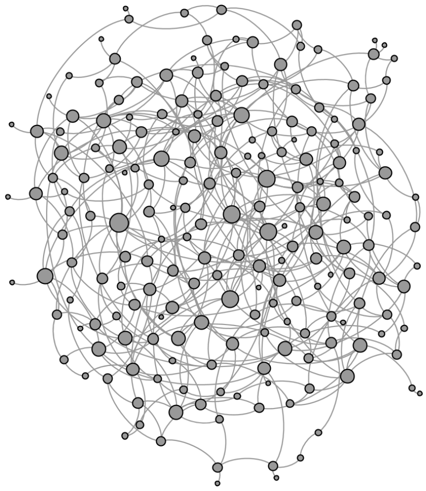


Figure B.1: A Random Graph.

Small-world Graph

The small-world phenomenon became famous after the works of Milgram in the sociological context [204] who found short chains of acquaintances connecting random people in the USA. More recently, the small-world characterization of graphs has been investigated by Watts and Strogatz [222, 223], who showed the presence of the small-world property in many types of networks such as actor acquaintances, the power grid, and neural networks in worms. It is obtained from a regular lattice that connects the nodes followed by a process of rewiring the edges with a certain probability $p \in [0, 1]$. The resulting graph has intermediate properties between the extreme situations of a regular lattice ($p = 0$) and a random graph ($p = 1$). In particular, small-world networks hold interesting properties: the characteristic path length is comparable to the one of a corresponding random graph ($L_{sw} \gtrsim L_{random}$), while the clustering coefficient has a value bigger than a random graph and closer to the one of a regular lattice ($CC_{sw} \gg CC_{random}$). A small-world graph with *order* 200 and *size* 399 is shown in Figure B.2.

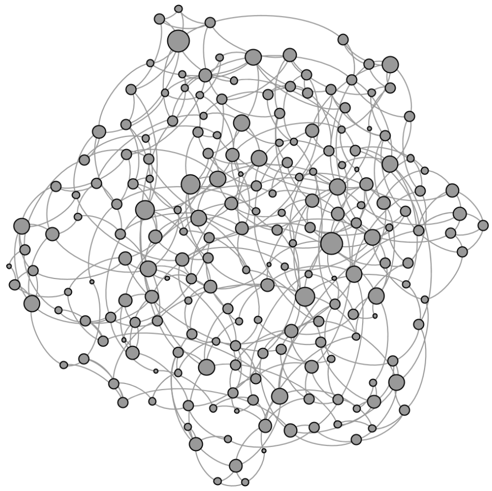


Figure B.2: A small-world graph.

Preferential Attachment

The preferential attachment model represents the phenomenon happening in real networks where a fraction of nodes has a high connectivity while the majority of nodes has small node degree. This model is built upon the observation by Barabási and Albert [15] of a typical pattern characterizing several type of natural and artificial networks. The basic idea is that whenever a node is added to the network and connects (through edges) to m other nodes, those with higher degree are preferred for connection. In other words, the probability to establish an edge with an existing node i is given by $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$ where k_i is the node degree of node i. One can see then that the more connected nodes have higher chances to acquire more and more edges over time in a sort of "rich gets richer" fashion; a phenomenon studied by Pareto [175] in relation to land ownership. The preferential attachment model reaches a stationary solution for the node degree probability that follows a power-law with $P(k) = \frac{2m^2}{k^3}$. A graph based on preferential attachment with *order* 200 and *size* 397 is shown in Figure B.3.

Figure B.3: A preferential attachment graph.

R-MAT

R-MAT (Recursive MATrix) is a model that exploits the representation of a graph through its adjacency matrix [40]. In particular, it applies a recursive method to create the adjacency matrix of a graph, thus obtaining a self-similar graph structure. This model captures the community-based pattern appearing in some real networks. Moreover, the generated graph is characterized by a power-law node degree distribution while showing a small diameter. The idea is to start with an empty $N \times N$ matrix and then divide the square matrix into four partitions in which the nodes are present with a certain probability for each partition, specifically probabilities a, b, c, d that sum to one. The procedure is then repeated dividing each partition again in four sub-partitions and associating the probabilities. The procedure stops when a 1×1 cell is reached in the iterative procedure. The a, b, c, d partitions of the adjacency matrix have particular meaning: a and d represent the portions containing nodes belonging to different communities, while b and c represent the nodes that act as a link for the different communities (e.g., in a social network people with interests both in topics mostly popular in either a or d community). The recursive nature of this algorithm creates a sort of sub communities at each round. A graph based on R-MAT model with *order* 222 and *size* 499 is shown in Figure B.4.

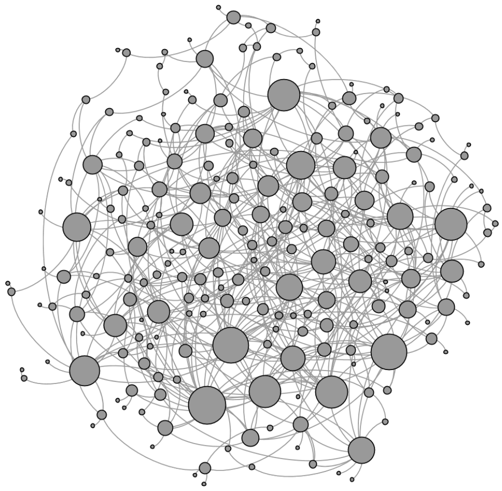


Figure B.4: A R-MAT graph.

Random Graph with Power-law

A Random Graph with power-law model generates networks characterized by a power-law in the node degree probability distribution ($P(k) \sim k^{-\gamma}$) having the majority of nodes with a low degree and a small amount of nodes with a very high degree. Power-law distributions are very common in many real life networks both created by natural processes (e.g., food-webs, protein interactions) and by artificial ones (e.g., airline travel routes, Internet routing, telephone call graphs) [13]. These are also referred to as *Scale-free networks* [16]. From the dynamic point of view, these networks are modeled by a preferential attachment model. In addition, reliability is a property of these graphs, that is, high degree of tolerance to random failures and high sensitivity to targeted attacks towards nodes with high degree or hubs [5, 138, 56].

This model is characterized by the exponent of the power-law (i.e., γ) which governs the degree of each node. The edges between the nodes are then wired in a random fashion. As we have shown earlier, the other way of constructing a graph that is compliant with a power-law based node degree distribution is through the growth of the network and preferential attachment based on node degree. A Random Graph with power-law with *order* 200 and *size* 399 is shown in Figure B.5.

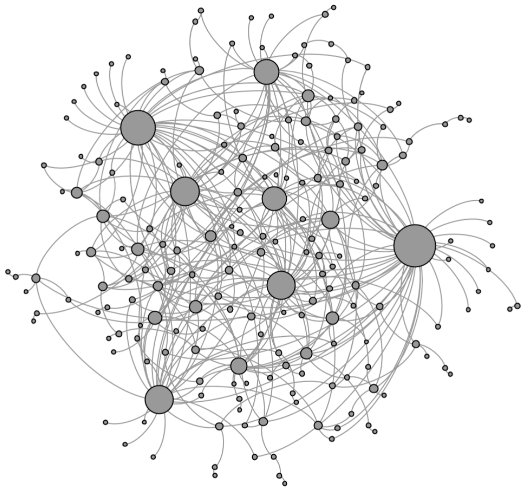


Figure B.5: A Random Graph with power-law graph.

Copying Model

Replicating the structure underlying the links of WWW pages brought to the Copying Model [107] which captures the tendency of members of communities with same interests to create pages on the web with a similar structure of links. The basic intuition is to select a node and a number (k) of edges to add to the node. Then with a certain probability β, the edges are linked independently and uniformly at random to k other nodes, while with probability $(1 - \beta)$ the k edges are copied from a randomly selected node u. If u has more than k edges, a subset is chosen, while if it has less than k edges they are anyway copied and the remaining are copied from another randomly chosen node. It leads to a distribution for the incoming degree that follows a power-law with a characteristic parameter $\gamma_{in} = \frac{1}{1-\beta}$. A graph based on Copying Model with *order* 200 and *size* 200 is shown in Figure B.6.

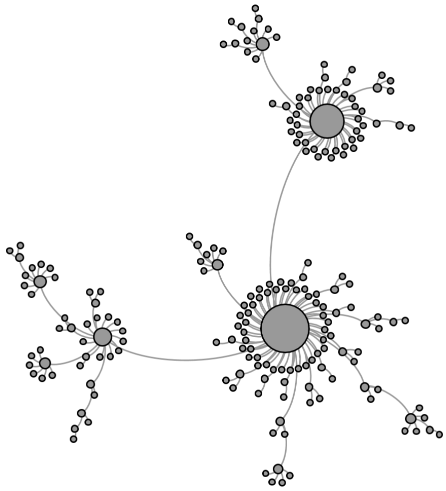


Figure B.6: A Copying Model graph.

Forest Fire

In order to capture dynamic aspects of the evolution of networks, Leskoveck *et al.* [118] proposed the Forest Fire model. The intuition is that networks tend to densify in connectivity and shrink in diameter (i.e., the greatest shortest path in the network) during the growth process. Technological, social and information networks exhibit this property. The model requires two parameters known as forward burning probability (p) and backward burning ratio (r). The graph grows over time and at each discrete time step a node v is added, then a node w, known as *ambassador*, is chosen at random between the other nodes of the graph and a link between v and w is added. A random number x (obtained from a binormal distribution with mean $(1-p)^{-1}$) is chosen and this is the number of out-links of node w that are selected. Then a fraction r times less than the out-links is chosen between the in-links and an edge is created with these as well. The process continues iterating choosing a new x number for each of the nodes v is now connected to. The idea, as the name of the model suggests, resembles the spreading of a fire in a forest that starts from the *ambassador* node to a fraction (based on the probability parameters) of nodes it is connected to and goes on in a sort of chain reaction. This model leads to heavy tails both in the distribution of in-degree and out-degree node degree. In addition, a power-law is shown in the densification process: a new coming node tends to have most of his links in the community of his *ambassador* and just few with other nodes. A graph based on Forest Fire model with *order* 200 and *size* 505 is shown in Figure B.7.

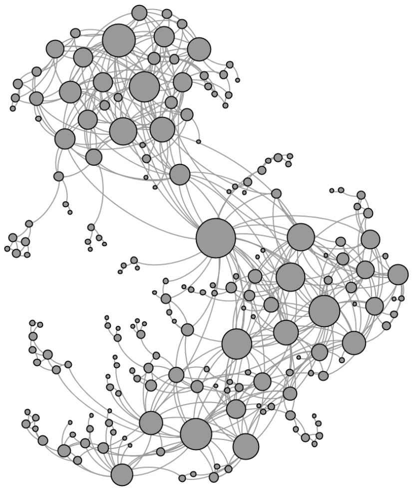


Figure B.7: A Forest Fire graph.

Kronecker Graph

A generating model with a recursive flavor similar to R-MAT uses the Kronecker product applied to the adjacency matrix of a graph [117]. The Kroneker product is a non conventional way of multiplying two matrices [80].

B.1. DEFINITION (KRONECKER PRODUCT). *Given two matrices A and B with dimension ($n \times m$) and ($n' \times m'$) the Kronecker product between A and B is a matrix C with dimension ($n \cdot n' \times m \cdot m'$) with the following structure:*

$$C = A \otimes B = \begin{pmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}B & a_{n,2}B & \cdots & a_{n,m}B \end{pmatrix}$$

B.2. DEFINITION (KRONECKER GRAPH). *Given two graphs G and H with adjacency matrices A(G) and A(H), a Kronecker graph is a graph whose adjacency matrix is obtained by the Kronecker product between the adjacency matrices of G and H.*

If the Kronecker product is applied to the same matrix, therefore multiplying the matrix with it elements recursively, a self-similar structure arises in the graph. This situation can be seen as the increase of a community in a network and the further differentiation in sub-communities while the network grows.

This model creates networks that show a densification in the connectivity with a shrinking diameter over time. The idea is to apply the Kronecker product to the same matrix recursively. The procedure to create a graph based on the Kronecker product starts with a $N \times N$ matrix where each x_{ij} element of the matrix represents a probability of having an edge between node i and j. Thereafter, at each time step the network grows so that at step k the network has N^k nodes. Applying the Kronecker product to the same matrix leads to the emergence of self-similar fractal-like structures at different scales. This structure mimics a quite natural process that is the recursive growth of communities inside communities that are a miniature copy of a big community (i.e., the whole graph structure) [117]. A Kronecker Graph with *order* 167 and *size* 267 is shown in Figure B.8.

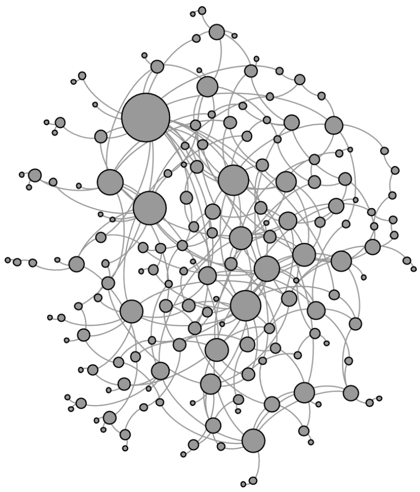


Figure B.8: A Kronecker graph.

B.2 Model Parameters for Synthetic Networks

In order to build graphs according to the models introduced in Section B.1 we used a set of specific parameters that are described below.

- **Random Graph**. For the $G_{N,M}$ model, the only parameters needed are the *order* and *size* of the graph to be generated. We use the values shown in Table 4.2 for the *order*, and the *size* is chosen accordingly to obtain an average node degree of two, four and six, respectively.

- **Small-world Graph**. In addition to *order*, the small-world model requires the specification of the average degree and the edge rewiring probability. For the first parameter, we simply provide a value to obtain the desired average node degree (i.e., $\langle k \rangle \approx 2$, $\langle k \rangle \approx 4$ and $\langle k \rangle \approx 6$). The latter parameter represents the probability of rewiring an edge, connecting a source node to a different destination node chosen at random. We choose an intermediate approach between the regular lattice (i.e., rewiring probability $p = 0$) and random graph extremes (i.e., rewiring probability $p = 1$). In fact, we choose a rewiring probability $p = 0.4$. This is to give slightly more emphasis to the regular structure of lattice than to the rewiring, since we expect the future grid to have more emphasis on a regular structure than random cabling. This last aspect also helps to satisfy the qualitative requirement of modularity.

- **Preferential Attachment**. For the creation of a graph based on growth and preferential attachment model of Barabási-Albert [15], the only parameters needed are the *order* and *size* of the graph to be generated. We use the values shown in Table 4.2 for the *order* parameter, while the *size* parameter is chosen accordingly to obtain an average node degree of two, four and six, respectively.

- **R-MAT**. The R-MAT model requires several parameters. First of all, *order* and *size* of the network, then the a, b, c, d parameters which represent the probabilities of the presence of an edge in a certain partition of the adjacency matrix. The order of the graph is chosen so that the nodes are a power of two, in particular 2^n where usually $n = \lceil log_2 N \rceil$. Therefore, we consider for this model the following values for the *order*: {32,128,256} for comparison with the low voltage, and {256,512,1024} for comparison with the medium voltage grids. For the probability parameters, since we have an undirected graph, we have $b = c$, in addition the ratio found between a and b, as in many real scenarios according to [40], is about 3:1. We assume a more highly connected community ($a = 0.46$) and a less connected community ($d = 0.22$) and a relative smaller connectivity between the two communities ($b = c = 0.16$).

- **Copying Model**. The Copying Model requires, in addition to the *order* of the graph, a value for the probability of copying (or not) edges from existing nodes. $(1 - \beta)$ is the probability of copying nodes from another node. In the present study, we fix $\beta = 0.2$ so as to have a high probability of having a direct (just one-hop since with probability 0.8 each new node copies the connections of another node and attaches directly to them) connection to what might be considered the most reliable energy sources present in the city or villages (at medium voltage level), while it represents single users or small aggregation of users with high energy capacity at low voltage level.

- **Forest Fire**. The Forest Fire model requires, in addition to the *order* of the graph, two values representing the probability of forward and backward spread of the "burning fire". We choose the same value for both probabilities since our graph is not directed. To avoid a flooding of edges, we choose few small values to assign to forward and backward probability ($p_{fwd} = p_{bwd} = 0.2$; $p_{fwd} = p_{bwd} = 0.3$; $p_{fwd} = p_{bwd} = 0.35$) that give realistic amounts of average edges incident to a node that can be compared with the models for which one is able to directly set *order* and *size*.

- **Random Graph with Power-law**. For the model representing Random Graph with power-law in node degree distribution, the parameters required are essentially the *order* of the network and the characteristic parameter of the power-law (known as the γ coefficient). For the first parameter, we use the usual dimensions (see Table 4.2), while for the latter some additional considerations are necessary. We test different types of power-law coefficients characterizing real technological networks. For the non-electrical technological networks (i.e., technological networks that are not power grids) we average the values of the power-law characteristic parameter described in [47]; the details of the parameters are shown in Table B.1. For the power grid networks the γ values represent:
  - the findings for the Western and Eastern high voltage U.S. power grid in [42]; the values are averaged to have a single γ, the details are shown in Table B.2;
  - the findings for the high voltage U.S. Western power grid in [15] that reports a value $\gamma = 4$;
  - the findings for the medium and low voltage Dutch grid that follow a power-law in [159]; the values are averaged to have a single γ, the details are shown in Table B.3.

Type of network	γ
Internet degree	2.12
Telephone calls received	2.09
Blackouts	2.3
Email address book size	3.5
Hits to web-sites	1.81
Links to web sites	2.336
Average	2.359

Table B.1: Power-law γ parameters for technological networks [47].

- **Kronecker Graph**. For the Kronecker model, the required parameter is the initial dimension of the square matrix to apply the Kronecker product: a 2 $\times$

Type of network	γ
Eastern Interconnection	3.04
Western System	3.09
Average	3.065

Table B.2: Power-law γ parameters for high voltage U.S. power grid [42].

Type of network	γ
LV#5	2.402
LV#10	1.494
MV#2	1.977
MV#3	2.282
Average	2.039

Table B.3: Power-law γ parameters for Dutch medium and low voltage power grid [159].

2 initiation matrix is a good starting model [117]. Once the structure of the matrix is defined, the initial parameters for the generation matrix need to be evaluated. With a 2×2 adjacency matrix for the initial graph G:

$$A(G) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

the parameters can be interpreted in a similar fashion as in R-MAT:

- a models the "core" part of the network and the tightness of its connectivity.
- d models the "perifery" part of the network and the connectivity inside it.
- b, c model the relationships and interconnections between the core and the periphery.

The findings of Leskovec *et al.* [117] applying the Kronecker modeling to many different networks report a common recurrent structure for the parameters of the 2×2 Kronecker matrix initiator. In particular, the parameters tend to follow the empirical rule $a \gg b \geqslant c \gg d$ and are usually $a \approx 1$, $b \approx c \approx 0.6$ and $d \approx 0.2$.

In this work, we consider two sets of parameters characterizing the Kronecker initiator matrix. The first set is extracted and averaged from the technological and social networks parameters extracted from real sample data [117]. The second set of parameters is obtained applying the fitting procedure to

Type of network	2×2 **Kronecker generator parameters**			
	a	c	b	d
Social-technological	0.9578	0.4617	0.4623	0.3162
power grid	0.4547	0.8276	0.8504	0.0186

Table B.4: Probability parameters for the 2×2 Kronecker matrix.

a Kronecker graph to the UCTE high voltage power grid data set used in [185, 192], the high voltage U.S. Western power grid data set used in [223], and the medium and low voltage samples data set used in [159]. All these values have been averaged to obtain just one 2×2 Kronecker generation matrix. A summary of the values for the parameters of the Kronecker matrix used is given in Table B.4. One notices a very different structure in the matrix parameters between the social and other diverse technological networks and the power grid.

Bibliography

[1] M. Aiello and S. Dustdar. A domotic infrastructure based on the web service stack. *Pervasive and Mobile Computing*, 4(4):506–525, 2008.

[2] M. Aiello, M. Papazoglou, J. Yang, M. Carman, M.Pistore, L. Serafini, and P. Traverso. A request language for web-services based on planning and constraint satisfaction. In *VLDB Ws on Tech. for E-Services (TES02)*, pages 76–85. Springer, 2002.

[3] W. Aiello, F. Chung, and L. Lu. Random evolution in massive graphs. In J. Abello, P. M. Pardalos, and M. G. C. Resende, editors, *Handbook of massive data sets*, pages 97–122. Kluwer Academic Publishers, Norwell, MA, USA, 2002.

[4] R. Albert, I. Albert, and G. Nakarado. Structural vulnerability of the north american power grid. *Physical Review E*, 69, 2004.

[5] R. Albert, H. Jeong, and A. L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[6] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21):11149–11152, 2000.

[7] C. W. Anderson, J. R. Santos, and Y. Y. Haimes. A Risk-based Input-Output Methodology for Measuring the Effects of the August 2003 Northeast Blackout. *Economic Systems Research*, 19(2):183–204, 2007.

[8] S. Arianos, E. Bompard, A. Carbone, and F. Xue. Power grid vulnerability: A complex network approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(1):013119, 2009.

[9] O. Asad, M. Erol-Kantarci, and H. Mouftah. Sensor network web services for demand-side energy management applications in the smart grid. In *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, pages 1176 –1180, 2011.

[10] K. Atkins, C. Barrett, and A. Marathe. A web services based artificial market. *2009 Winter Simulation Conf. (WSC)*, pages 3047–3054, 2009.

[11] P. Balducci, J. Roop, L. Schienbein, J. DeSteese, and M. Weimar. Electrical power interruption cost estimates for individual industries, sectors, and US economy. Technical report, U.S. Department of Energy, 2002.

[12] A. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3–4):590 – 614, 2002.

[13] A. L. Barabási. *Linked: How Everything Is Connected to Everything Else and What It Means*. Plume, reissue edition, 2003.

[14] A. L. Barabási. Scale-free networks: a decade and beyond. *Science*, 325(5939):412–3, 2009.

[15] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509, 1999.

[16] A. L. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the World Wide Web. *Physica A: Statistical Mechanics and its Applications*, 281(1-4):69–77, 2000.

[17] M. Baran and F. Wu. Network reconfiguration in distribution systems for loss reduction and load balancing. *Power Delivery, IEEE Transactions on*, 4(2):1401 –1407, 1989.

[18] D. Becker, H. Falk, J. Gillerman, S. Mauser, R. Podmore, and L. Schneberger. Standards-based approach integrates utility applications. *IEEE Computer Applications in Power*, 13(4):13–20, 2000.

[19] H. Beitollahi and G. Deconinck. Peer-to-peer networks applied to power grid. In *International conference on Risks and Security of Internet and Systems (CRiSIS)*, 2007.

[20] D. Berardi, D. Calvanese, G. D. Giacomo, and M. Mecella. Reasoning about Actions for e-Service Composition. In *ICAPS'03 Ws on Planning for Web Services*, 2003.

[21] L. M. A. Bettencourt. The origins of scaling in cities. *Science*, 340(6139):1438–1441, 2013.

[22] F. Bliek, A. van den Noort, B. Roossien, R. Kamphuis, J. de Wit, J. van der Velde, and M. Eijgelaar. Powermatching city, a living lab smart grid demonstration. In *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, pages 1–8, 2010.

[23] J. Bolderdijk, L. Steg, E. Geller, P. Lehman, and T. Postmes. Comparing the effectiveness of monetary versus moral motives in environmental campaigning. *Nature Climate Change*, 2012.

[24] B. Bollobas. *Graph theory : an introductory course*. Springer Verlag,, New York, 1979.

[25] B. Bollobas. *Modern Graph Theory*. Springer, 1998.

[26] E. Bompard, R. Napoli, and F. Xue. Analysis of structural vulnerabilities in power transmission grids. *International Journal of Critical Infrastructure Protection*, 2(1-2):5–12, 2009.

[27] E. Bompard, R. Napoli, and F. Xue. Social and cyber factors interacting over the infrastructures: A mas framework for security analysis. In R. R. Negenborn, Z. Lukszo, and H. Hellendoorn, editors, *Intelligent Infrastructures*, volume 42 of *Intelligent Systems, Control and Automation: Science and Engineering*, pages 211–234. Springer Netherlands, 2010.

[28] E. Bompard, D. Wu, and F. Xue. The concept of betweenness in the analysis of power grid vulnerability. In *Complexity in Engineering, 2010.*, pages 52 –54, 2010.

[29] F. Bouffard. The challenge with building a business case for smart grids. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1 –3, 2010.

[30] P. Bresesti, M. Gallanti, and D. Lucarella. Market-based generation and transmission expansions in the competitive market. In *Power Engineering Society General Meeting, 2003, IEEE*, volume 1, page 4 vol. 2666, 2003.

[31] R. Brown. Impact of smart grid on distribution system design. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1 –4, 2008.

[32] C. D. Brummitt, R. M. DSouza, and E. A. Leicht. Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences*, 109(12):E680–E689, 2012.

[33] N. Capodieci, E. Alsina, and G. Cabri. A context-aware agent-based approach for deregulated energy market. In *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2012 IEEE 21st International Workshop on*, pages 16 –21, 2012.

[34] N. Capodieci, G. Cabri, G. A. Pagani, and M. Aiello. Adaptive game-based agent negotiation in deregulated energy markets. In *Collaboration Technologies and Systems (CTS), 2012 International Conference on*, pages 300–307, 2012.

[35] N. Capodieci, G. Cabri, G. A. Pagani, and M. Aiello. An agent-based application to enable deregulated energy markets. In *Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th Annual*, pages 638–647, 2012.

[36] N. Capodieci, G. A. Pagani, G. Cabri, and M. Aiello. Smart meter aware domestic energy trading agents. In *Proceedings of the 2011 workshop on E-energy market challenge*, IEEMC '11, pages 1–10, New York, NY, USA, 2011. ACM.

[37] F. Casati, M. Sayal, and M.-C. Shan. Developing e-services for composing e-services. In *Conf. on Advanced Inf. Sys. Engineering (CAiSE)*, pages 171–186. Springer, 2001.

[38] V. Chaitanya, M. Reddy, and D. Mohanta. Topological analysis of eastern region of indian power grid. In *Environment and Electrical Engineering (EEEIC), 2011 10th International Conference on*, pages 1 –4, 2011.

[39] D. Chakrabarti and C. Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM Computing Surveys (CSUR)*, 38(1):2, 2006.

[40] D. Chakrabarti, Y. Zhan, and C. Faloutsos. R-MAT: A Recursive Model for Graph Mining. In *Fourth SIAM International Conference on Data Mining*, 2004.

[41] D. Chappell. *Enterprise Service Bus*. O'Reilly, 2004.

[42] D. P. Chassin and C. Posse. Evaluating North American electric grid reliability using the Barabási Albert network model. *Physica A: Statistical Mechanics and its Applications*, 355:667–677, 2005.

[43] P. Chiradeja. Benefit of Distributed Generation: A Line Loss Reduction Analysis. *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, pages 1–5, 2005.

[44] K. Chmiel, M. Gawinecki, P. Kaczmarek, M. Szymczak, and M. Paprzycki. Efficiency of JADE agent platform. *Scientific Programming*, 13(2):159–172, 2005.

[45] J. Choi, T. Tran, A. El-Keib, R. Thomas, H. Oh, and R. Billinton. A method for transmission system expansion planning considering probabilistic reliability criteria. *Power Systems, IEEE Transactions on*, 20(3):1606 – 1615, 2005.

[46] S. Civanlar, J. Grainger, H. Yin, and S. Lee. Distribution feeder reconfiguration for loss reduction. *Power Delivery, IEEE Transactions on*, 3(3):1217 –1223, 1988.

[47] A. Clauset, C. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[48] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Physical review letters*, 85(21):4626–8, 2000.

[49] R. Cohen and S. Havlin. *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, 2010.

[50] T. Considine. Ontological requirements of the service oriented grid. In *Grid-Interop - The road to an interoperable grid, Atlanta, Georgia, USA*, 2008.

[51] R. Cossent, T. Gómez, and P. Frías. Towards a future with large penetration of distributed generation: Is the current regulation of electricity distribution ready? regulatory recommendations under a european perspective. *Energy Policy*, 37(3):1145–1155, 2009.

[52] T. Cover and P. Hart. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1):21 –27, 1967.

[53] W. T. Cox and T. Considine. Architecturally significant interfaces for the smart grid. In *Grid-Interop - The road to an interoperable grid, Denver, Colorado, USA*, 2009.

[54] D. Crawford and J. Holt, S.B. A mathematical optimization technique for locating and sizing distribution substations, and deriving their optimal service areas. *Power Apparatus and Systems, IEEE Transactions on*, 94(2):230 – 235, 1975.

[55] B. Croxford and K. Scott. Can pv or solar thermal systems be cost effective ways of reducing co 2 emissions for residential buildings? In *Solar 2006: Renewable Energy - Key to Climate Recovery*, American Solar Energy Society, 2006.

[56] P. Crucitti, V. Latora, and M. Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4):3–6, 2004.

[57] P. Crucitti, V. Latora, and M. Marchiori. A topological analysis of the italian electric power grid. *Physica A: Statistical Mechanics and its Applications*, 338(1-2):92 – 97, 2004.

[58] P. Crucitti, V. Latora, and M. Marchiori. Locating critical lines in high-voltage electrical power grids. *Fluctuation and Noise Letters*, 5(2):L201–L208, 2005.

[59] F. Dahlfors and J. Pilling. Integrated information systems in a privatized and deregulated electri city market. In *1995 Int. Conf. on Energy Management and Power Delivery EMPD '95*, pages 249–254. IEEE, 1995.

[60] J. Ding, X. Bai, W. Zhao, Z. Fang, Z. Li, and M. Liu. The improvement of the small-world network model and its application research in bulk power system. In *Power System Technology, 2006. International Conference on*, pages 1–5, 2006.

[61] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, 2002.

[62] S. Dunn. Micropower: The next electrical era. *Worldwatch Paper*, 2000.

[63] S. Dustdar and W. Schreiner. A survey on web services composition. *Int. J. Web Grid Serv.*, 1(1):1–30, 2005.

[64] A. Dwivedi, X. Yu, and P. Sokolowski. Identifying vulnerable lines in a power network using complex network theory. In *Industrial Electronics, 2009. IEEE International Symposium on*, pages 18–23, 2009.

[65] A. Dwivedi, X. Yu, and P. Sokolowski. Analyzing power network vulnerability with maximum flow based centrality approach. In *Industrial Informatics (INDIN), 2010 8th IEEE International Conference on*, pages 336 –341, 2010.

[66] J. Ekanayake, N. Jenkins, K. Liyanage, J. Wu, and A. Yokoyama. *Smart Grid: Technology and Applications*. Wiley, 2012.

[67] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[68] S. K. Ethiraj and D. Levinthal. Modularity and innovation in complex systems. *Management Science*, 50(2):pp. 159–173, 2004.

[69] J. Farago, A. Greenwald, and K. Hall. Fair and efficient solutions to the santa fe bar problem. In *In Proceedings of the Grace Hopper Celebration of Women in Computing 2002*, 2002.

[70] K. Fekete, S. Nikolovski, D. Puzak, G. Slipac, and H. Keko. Agent-based modelling application possibilities for croatian electricity market simulation. In *Electricity Market, 2008. EEM 2008. 5th International Conference on European*, pages 1–6, 2008.

[71] A. Fratkin and L. Ostrov. Ehv transmission grid design in israel electric corporation's power system: approaches and experience. In *Electrical and Electronics Engineers in Israel, 1996., Nineteenth Convention of*, pages 475 –478, 1996.

[72] L. Freris and D. Infield. *Renewable Energy in Power Systems*. Wiley, 2008.

[73] T. L. Friedman. *Hot, Flat, and Crowded: Why We Need a Green Revolution - and How It Can Renew America*. Picador, 2008.

[74] L. Garver. Transmission network estimation using linear programming. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(7):1688 –1697, 1970.

[75] C. W. Gellings. A consumer portal at the junction of electricity, communications, and consumer energy services. *The Electricity Journal*, 17(9):78 – 84, 2004.

[76] I. Georgievski, V. Degeler, G. A. Pagani, T. A. Nguyen, A. Lazovik, and M. Aiello. Optimizing energy costs for offices connected to the smart grid. *Smart Grid, IEEE Transactions on*, 3(4):2273–2285, 2012.

[77] J. K. Gershenson, G. J. Prasad, and Y. Zhang. Product modularity: definitions and benefits. *Journal of Engineering Design*, 14(3):295–313, 2003.

[78] H. Ghenniwa. Web-services infrastructure for information integration in power systems. In *IEEE Power Engineering Society*, page 8 pp. IEEE, 2006.

[79] G. Gilbert, Y. Chow, D. Bouchard, and M. Salama. Optimization of high voltage substations using a random walk technique. In *Transmission and Distribution Construction, Operation and Live-Line Maintenance (ESMO), 2011 IEEE PES 12th International Conference on*, pages 1 –7, 2011.

[80] A. Graham. *Kronecker products and matrix calculus: with applications*. Ellis Horwood series in mathematics and its applications. Horwood, 1981.

[81] L. L. Grigsby, editor. *The Electric Power Engineering Handbook, Second Edition*. CRC Press, 2007.

[82] Z. Guohua, W. Ce, Z. Jianhua, Y. Jingyan, Z. Yin, and D. Manyin. Vulnerability assessment of bulk power grid based on complex network theory. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. Third International Conference on*, pages 1554–1558, 2008.

[83] P. Han and M. Ding. Analysis of Cascading Failures in Small-world Power Grid. *International Journal of Energy Science*, 1(2):99–104, 2011.

[84] C. Harris. *Electricity Markets: Pricing, Structures and Economics*. Wiley, 2006.

[85] J. A. Harris and F. G. Benedict. A biometric study of human basal metabolism. *Proceedings of the National Academy of Sciences*, 4(12):370–373, 1918.

[86] P. Hines and S. Blumsack. A centrality measure for electrical networks. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, page 185, 2008.

[87] P. Hines, E. Cotilla-Sanchez, and S. Blumsack. Do topological models provide good information about electricity infrastructure vulnerability? *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(3):033122, 2010.

[88] T. E. Hoff. Using distributed resources to manage risks caused by demand uncertainty. *The Energy Journal*, 18(Special I):63–84, 1997.

[89] A. J. Holmgren. Using Graph Models to Analyze the Vulnerability of Electric Power Networks. *Risk Analysis*, 26(4):955–969, 2006.

[90] IEA. Key world energy statistics 2013. *International Energy Agency*, 2013.

[91] N. R. Jennings. On agent-based software engineering. *Artificial Intelligence*, 117(2):277 – 296, 2000.

[92] Y. Jia-hai, Y. Shun-kun, and H. Zhao-guang. A multi-agent trading platform for electricity contract market. *2005 International Power Engineering Conf.*, 102206:1024–1029 Vol. 2, 2005.

[93] P. L. Joskow. Lessons learned from electricity market liberalization. *The Energy Journal*, 29(Special I):9–42, 2008.

[94] S. Karnouskos. The cooperative internet of things enabled smart grid. In *Proceedings of the 14th IEEE International Symposium on Consumer Electronics (ISCE2010), June 07-10, Braunschweig, Germany*, 2010.

[95] S. Karnouskos. Future smart grid prosumer services. In *Innovative Smart Grid Technologies (ISGT Europe), 2011 2nd IEEE PES International Conference and Exhibition on*, pages 1 –2, 2011.

[96] S. Karnouskos, P. Da Silva, and D. Ilic. Energy services for the smart grid city. In *Digital Ecosystems Technologies (DEST), 2012 6th IEEE International Conference on*, pages 1 –6, 2012.

[97] S. Karnouskos and T. N. D. Holanda. Simulation of a Smart Grid City with Software Agents. *2009 Third UKSim European Symposium on Computer Modeling and Simulation*, pages 424–429, 2009.

[98] S. Karnouskos, O. Terzidis, and P. Karnouskos. An advanced metering infrastructure for future energy networks. In *IFIP/IEEE 1st International Conference on New Technologies, Mobility and Security (NTMS 2007), Paris, France*, pages 597–606, 2007.

[99] S. Karnouskos, A. Weidlich, K. Kok, C. Warmer, J. Ringelstein, P. Selzam, A. Dimeas, and S. Drenkard. Field trials towards integrating smart houses with the smart grid. In *Energy-Efficient Computing and Networking*, volume 54 of *LNCS*, pages 114–123. Springer, 2011.

[100] R. L. Keeney. Decision analysis: An overview. *Operations Research*, 30(5):pp. 803–838, 1982.

[101] W. Kehe, S. Qianhui, D. Cheng, and Y. Haiyani. The research of power grid data integration and sharing platform based on soa. *Computer Science and Engineering, 2009. WCSE '09. Second Int. Ws on*, 1:106 –109, 2009.

[102] M. Kendall. *Rank correlation methods*. Griffin, London, 1948.

[103] A. Khan and H. Mouftah. Web services for indoor energy management in a smart grid environment. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*, pages 1036 –1040, 2011.

[104] T. Khoa, P. Binh, and H. Tran. Optimizing Location and Sizing of Distributed Generation in Distribution Systems. *2006 IEEE PES Power Systems Conference and Exposition*, pages 725–732, 2006.

[105] C. J. Kim and O. B. Obah. Vulnerability assessment of power grid using graph topological indices. *International Journal of Emerging Electric Power Systems*, 8(6):1–15, 2007.

[106] D. King, W. Boyson, and J. Kratochvil. Analysis of factors influencing the annual energy production of photovoltaic systems. In *Photovoltaic Specialists Conference, 2002. Conference Record of the Twenty-Ninth IEEE*, pages 1356 – 1361, 2002.

[107] J. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The web as a graph: Measurements, models, and methods. In *Proceedings of the 5th annual international conference on Computing and combinatorics*, pages 1–17. Springer-Verlag, 1999.

[108] J. M. Kleinberg. The small-world phenomenon: an algorithm perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, STOC '00, pages 163–170, New York, NY, USA, 2000. ACM.

[109] P. D. Klemperer and M. A. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):pp. 1243–1277, 1989.

[110] K. Kok, Z. Derzsi, M. Hommelberg, C. Warmer, R. Kamphuis, and H. Akkermans. Agent-based electricity balancing with distributed energy resources, a multiperspective case study. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, page 173, 2008.

[111] K. Kok and G. Venekamp. Market-based control in decentralized electrical power systems. In *First International Workshop on Agent Technologies for Energy Systems, ATES2010, Toronto*, 2010.

[112] K. Kok, C. J. Warmer, and I. G. Kamphuis. Powermatcher: multiagent control in the electricity infrastructure. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, AAMAS '05, pages 75–82, New York, NY, USA, 2005. ACM.

[113] L. L. Lai, T. Motshegwa, H. Suasinghe, N. Rajkumar, and R. Blach. Feasibility study with agents on energy trading. In *Advances in Power System Control, Operation and Management, 2000. APSCOM-00. 2000 International Conference on*, volume 2, pages 505–510 vol.2, 2000.

[114] A. Lazovik, M. Aiello, and M. Papazoglou. Planning and monitoring the execution of web service requests. *Journal on Digital Libraries*, 6(3):235–246, 2006.

[115] A. D. Le, M. Kashem, M. Negnevitsky, and G. Ledwich. Optimal Distributed Generation Parameters for Reducing Losses with Economic Consideration. *2007 IEEE Power Engineering Society General Meeting*, pages 1–8, 2007.

[116] S. Lee, K. Hicks, and E. Hnyilicza. Transmission expansion by branch-and-bound integer programming with optimal cost - capacity curves. *Power Apparatus and Systems, IEEE Transactions on*, PAS-93(5):1390 –1400, 1974.

[117] J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. Kronecker graphs: An approach to modeling networks. *The Journal of Machine Learning Research*, 11:985–1042, 2010.

[118] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, KDD '05, pages 177–187, New York, NY, USA, 2005. ACM.

[119] F. Leuthold, H. Weigt, and C. von Hirschhausen. Efficient pricing for european electricity networks - the theory of nodal pricing applied to feeding-in wind in germany. *Utilities Policy*, 16(4):284 – 291, 2008. European Regulatory Perspectives.

[120] F. Leymann, D. Roller, and M.-T. Schmidt. Web services and business process management. *IBM Systems Journal*, 41(2):198–211, 2002.

[121] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.*, 58(7):1019–1031, 2007.

[122] R. Lincoln, S. Galloway, and G. Burt. Open source, agent-based energy market simulation with python. In *Energy Market, 2009. EEM 2009. 6th International Conference on the European*, pages 1–5, 2009.

[123] A. B. Lovins, E. K. Datta, T. Feiler, K. R. Rabago, J. N. Swisher, A. Lehmann, and K. Wicker. *Small is profitable: the hidden economic benefits of making electrical resources the right size*. Rocky Mountain Institute, 2002.

[124] G. Lu, D. De, and W.-Z. Song. Smartgridlab: A laboratory-based smart grid testbed. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 143 –148, 2010.

[125] L. Lü and T. Zhou. Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and its Applications*, 390(6):1150 – 1170, 2011.

[126] A. Machias, E. Dialynas, and C. Protopapas. An expert system approach to designing and testing substation grounding grids. *Power Delivery, IEEE Transactions on*, 4(1):234 – 240, 1989.

[127] R. Mackiewicz. The Benefits of Standardized Web Services Based on the IEC 61970 Generic Interface Definition. In *IEEE Power Systems Conf. and Expo.*, pages 491–494, 2006.

[128] M. Marmiroli and H. Suzuki. Web-based framework for electricity market. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2000. Proceedings. DRPT 2000. International Conference on*, pages 471–475, 2000.

[129] C. Marnay and G. Venkataramanan. Microgrids in the evolving electricity generation and delivery infrastructure. In *In IEEE Power Engineering Society General Meeting*. IEEE, 2006.

[130] J. Massey, Frank J. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):pp. 68–78, 1951.

[131] F. McDonald. *Insull*. Beard Books, 1962.

[132] S. Mei, X. Zhang, and M. Cao. *Power Grid Complexity*. Springer, 2011.

[133] A. Mercurio, A. Di Giorgio, and P. Cioci. Open-source implementation of monitoring and controlling services for emsscada systems by means of web services - iec 61850 and iec 61970 standards. *Power Delivery, IEEE Transactions on*, 24(3):1148–1153, 2009.

[134] A. Merlin and H. Back. Search for a minimal-loss operating spanning tree configuration in an urban power distribution system. In *Proc. of the Fifth Power System Conference (PSCC), Cambridge*, pages 1–18, 1975.

[135] S. Milgram. The small world problem. *Psychology Today*, 2:60–67, 1967.

[136] M. Mitchell. *Complexity - A Guided Tour*. Oxford University Press, 2009.

[137] D. Moneta, L. Bisone, G. Mauri, and R. Meda. New interactions between lv customers and the network: further possibilities for home automation functions. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 2886–2891, 2007.

[138] Y. Moreno, R. Pastor-Satorras, A. Vazquez, and A. Vespignani. Critical load and congestion instabilities in scale-free networks. *Europhysics Letters*, 62, 2003.

[139] M. G. Morgan, J. Apt, L. B. Lave, M. D. Ilic, M. Sirbu, and J. M. Peha. The many meanings of "smart grid". Technical report, Carnegie Mellon University, 2009.

[140] K. Moslehi, A. Kumar, E. Dehdashti, P. Hirsch, and W. Wu. Distributed autonomous real-time system for power system operations - a conceptual overview. In *IEEE PES Power Systems Conf. and Expo*, volume 1pp, pages 1705–1712. IEEE, 2004.

[141] M. Munasinghe. Engineering-economic analysis of electric power systems. *Proceedings of the IEEE*, 72(4):424 – 461, 1984.

[142] B. Murray. *Power Markets and Economics: Energy Costs, Trading, Emissions*. Wiley, 2009.

[143] A. Nasiruzzaman and H. Pota. Critical node identification of smart power system using complex network framework based centrality approach. In *North American Power Symposium, 2011*, pages 1 –6, 2011.

[144] S. Nath and R. Somnath. Network Reconfiguration for Electrical Loss Minimization. *Inernational Journal of Instrumentation, Control and Automation*, 1, 2011.

[145] National Energy Technology Laboratory. A system view of the modern grid. Technical report, U.S. Department of Energy - Office of Electricity Delivery and Energy Reliability, 2007.

[146] National Energy Technology Laboratory. A vision for the modern grid. Technical report, U.S. Department of Energy - Office of Electricity Delivery and Energy Reliability, 2007.

[147] National Energy Technology Laboratory. Building a smart grid business case. Technical report, U.S. Department of Energy - Office of Electricity Delivery and Energy Reliability, 2009.

[148] National Grid. Undergrounding high voltage electricity transmission - the technical issues. Technical report, National Grid, 2009.

[149] Netbeheer Nederland. Betrouwbaarheid van elektriciteitsnetten in nederland - resultaten 2012. Technical Report RM-ME-13L10440006, Netbeheer Nederland, 2013.

[150] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, 2002.

[151] M. E. J. Newman. The structure and function of complex networks. *SIAM REVIEW*, 45:167–256, 2003.

[152] M. E. J. Newman. Analysis of weighted networks. *Phys. Rev. E*, 70(5):056131, 2004.

[153] M. E. J. Newman. *Networks: An Introduction*. OUP Oxford, 2010.

[154] OASIS EMIX Technical Committee. Energy market information exchange (emix). Technical Report Version 1.0, 2012.

[155] P. D. O'Brien and R. C. Nicol. Fipa towards a standard for software agents. *BT Technology Journal*, 16(3):51–59, 1998.

[156] Office of the National Coordinator for Smart Grid Interoperability. Nist framework and roadmap for smart grid interoperability standards, release 1.0. Technical Report NIST Special Publication 1108, National Institute of Standards and Technology, 2010.

[157] G. A. Pagani and M. Aiello. Energy market trading systems in G6 countries. Technical Report JBI preprint 2010-6-01, JBI, University of Groningen, 2010. Available at `http://www.cs.rug.nl/~andrea/publications/energyMarketG6.pdf`.

[158] G. A. Pagani and M. Aiello. Towards a service-oriented energy market: Current state and trend. In E. Maximilien, G. Rossi, S.-T. Yuan, H. Ludwig, and M. Fantinato, editors, *Service-Oriented Computing*, volume 6568 of *Lecture Notes in Computer Science*, pages 203–209. Springer Berlin Heidelberg, 2011.

[159] G. A. Pagani and M. Aiello. Towards decentralization: A topological investigation of the medium and low voltage grids. *Smart Grid, IEEE Transactions on*, 2(3):538 –547, 2011.

[160] G. A. Pagani and M. Aiello. Power grid network evolutions for local energy trading. Technical Report Available at arXiv:1201.0962, JBI, University of Groningen, 2012.

[161] G. A. Pagani and M. Aiello. Service orientation and the smart grid state and trends. *Service Oriented Computing and Applications*, 6(3):267–282, 2012.

[162] G. A. Pagani and M. Aiello. A complex network approach for identifying vulnerabilities of the medium and low voltage grid. *International Journal of Critical Infrastructures*, (To appear), 2013.

[163] G. A. Pagani and M. Aiello. Cost and benefits of denser topologies for the smart grid. In E. Gelenbe and R. Lent, editors, *Computer and Information Sciences III*, pages 73–81. Springer London, 2013.

[164] G. A. Pagani and M. Aiello. From the grid to the smart grid, topologically. Technical Report Available at arXiv:1305.0458, JBI, Univ. of Groningen, 2013.

[165] G. A. Pagani and M. Aiello. Generating realistic dynamic prices and services for the smart grid. Technical Report JBI preprint 2013-12-01, JBI, University of Groningen, 2013. Available at `http://www.cs.rug.nl/~andrea/publications/techRepBB.pdf`.

[166] G. A. Pagani and M. Aiello. Modeling the last mile of the smart grid. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pages 1–6, 2013.

[167] G. A. Pagani and M. Aiello. The power grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Applications*, 392(11):2688 – 2700, 2013.

[168] G. A. Pagani and M. Aiello. Power grid complex network evolutions for the smart grid. *Physica A: Statistical Mechanics and its Applications*, 396(0):248 – 266, 2014.

[169] S. Pahwa, A. Hodges, C. Scoglio, and S. Wood. Topological analysis of the power grid and mitigation strategies against cascading failures. In *Systems Conference, 2010 4th Annual IEEE*, pages 272–276, 2010.

[170] P. Paiva, H. Khodr, J. Dominguez-Navarro, J. Yusta, and A. Urdaneta. Integral planning of primary-secondary distribution systems using mixed integer linear programming. *Power Systems, IEEE Transactions on*, 20(2):1134 – 1143, 2005.

[171] M. Panahi, W. Nie, and K.-J. Lin. A Framework for Real-Time Service-Oriented Architecture. *IEEE Conf. on Commerce and Enterprise Computing*, pages 460–67, 2009.

[172] M. P. Papazoglou. *Web Services: Principles and Technology*. Pearson, 2007.

[173] M. P. Papazoglou and D. Georgakopoulos. Service-oriented computing. *Commun. ACM*, 46(10):24–28, 2003.

[174] M. P. Papazoglou and W.-j. Van Den. Service oriented architectures : approaches , technologies and researc h issues. *Vldb Journal*, pages 389–415, 2007.

[175] V. Pareto. *Manual of political economy (manuale di economia politica)*. Kelley, New York, 1971 (1906). Translated by Ann S. Schwier and Alfred N. Page.

[176] D. Pepyne. Topology and cascading line outages in power grids. *Journal of Systems Science and Systems Engineering*, 16:202–221, 2007.

[177] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya. Local assortativity and growth of Internet. *The European Physical Journal B*, 70(2):275–285–285, 2009.

[178] M. Postina, S. Rohjans, U. Steffens, and M. Uslar. Views on service oriented architectures in the context of smart grids. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 25 –30, 2010.

[179] R. Kinney, P. Crucitti, R. Albert, and V. Latora. Modeling cascading failures in the north american power grid. *European Physical Journal B*, 46(1):101–107, 2005.

[180] L. Ramesh, S. P. Chowdhury, S. Chowdhury, A. A. Natarajan, and C. T. Gaunt. Minimization of power loss in distribution networks by different techniques. *Inernational Journal of Electrical Power and Energy Systems Engineering*, 3, 2009.

[181] S. Rehman, M. A. Bader, and S. A. Al-Moallem. Cost of solar energy generated using pv panels. *Renewable and Sustainable Energy Reviews*, 11(8):1843 – 1857, 2007.

[182] C. Roe, S. Meliopoulos, R. Entriken, and S. Chhaya. Simulated demand response of a residential energy management system. In *Energytech, 2011 IEEE*, pages 1 –6, 2011.

[183] M. Rosas-Casals. Power grids as complex networks: Topology and fragility. In *Complexity in Engineering, 2010.*, pages 21 –26, 2010.

[184] M. Rosas-Casals and B. Corominas-Murtra. Assessing European power grid reliability by means of topological measures. *WIT transactions on ecology and the environment*, 121:527–537, 2009.

[185] M. Rosas-Casals, S. Valverde, and R. V. Solé. Topological vulnerability of the European power grid under errors and attacks. *International Journal of Bifurcation and Chaos*, 17(07):2465, 2007.

[186] V. Rosato, S. Bologna, and F. Tiriticco. Topological properties of high-voltage electrical transmission networks. *Electric Power Systems Research*, 77(2):99–105, 2007.

[187] L. Schleisner. Life cycle assessment of a wind farm and related externalities. *Renewable Energy*, 20(3):279 – 288, 2000.

[188] A. Schuelke and K. Erickson. Serving solar variations with consumption control of smart appliances and electric vehicles. In *Innovative Smart Grid Technologies (ISGT Europe), 2011 2nd IEEE PES International Conference and Exhibition on*, pages 1 –8, 2011.

[189] G. Shakhnarovich, T. Darrell, and P. Indyk. *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)*. The MIT Press, 2006.

[190] H. Shariati, H. Askarian Abyaneh, M. Javidi, and F. Razavi. Transmission expansion planning considering security cost under market environment. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, pages 1430 –1435, 2008.

[191] M. H. Shariatkhah and M. R. Haghifam. Using feeder reconfiguration for congestion management of smart distribution network with high dg penetration. In *Integration of Renewables into the Distribution Grid, CIRED 2012 Workshop*, pages 1 –4, 2012.

[192] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde. Robustness of the European power grids under intentional attack. *Physical Review E*, 77(2):1–7, 2008.

[193] Y. Song, G. Tang, Z. Yang, and J. Hu. The Technical Implementation of the Electricity Market Operation Syste m. In *2005 IEEE/PES Transmission and Distribution Conf. and Exhibition: Asia and Pacific*, pages 1–6. IEEE, 2005.

[194] P. Sotkiewicz and J. Vignolo. Nodal pricing for distribution networks: efficient pricing for efficiency enhancing dg. *Power Systems, IEEE Transactions on*, 21(2):1013–1014, 2006.

[195] P. Stephenson. Electricity market trading. *Power Engineering J.*, 15(6):277, 2001.

[196] M. Strobbe, T. Verschueren, K. Mets, S. Melis, C. Develder, F. De Turck, T. Pollet, and S. Van de Veire. Design and evaluation of an architecture for future smart grid service provisioning. In *Network Operations and Management Symposium (NOMS), 2012 IEEE*, pages 1203 –1206, 2012.

[197] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–76, 2001.

[198] S. Sucic, A. Martinic, and A. Kekelj. Utilizing standards-based semantic services for modeling novel smart grid supervision and remote control frameworks. In *Industrial Technology (ICIT), 2012 IEEE International Conference on*, pages 409 –414, 2012.

[199] K. Sun. Complex networks theory: A new method of research in power grid. In *Transmission and Distribution Conference and Exhibition: Asia and Pacific*, pages 1–6, 2005.

[200] D. T. Swift-Hook. Grid-connected intermittent renewables are the last to be stored. *Renewable Energy*, 35(9):1967 – 1969, 2010.

[201] H. Takamori, K. Nagasaka, and E. Go. Toward Designing Value Supportive Infrastructure for Electricity Trading. *The 9th IEEE Int. Conf. on E-Commerce Technology (CEC-EEE 2007)*, pages 167–174, 2007.

[202] H. Takamori, K. Nagasaka, and E. Go. Toward designing value supportive infrastructure for electricity trading. In *E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007. CEC/EEE 2007.*, pages 167 –174, 2007.

[203] W. Taqqali and N. Abdulaziz. Smart grid and demand response technology. In *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, pages 710 –715, 2010.

[204] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.

[205] W. Tsai, Y.-h. Lee, Z. Cao, Y. Chen, and B. Xia o. RTSOA: Real-Time Service-Oriented Architecture. *2006 Second IEEE Int. Sym. Service-Oriented System Engineering (SOSE'06)*, pages 49–56, 2006.

[206] L. Tsoukalas and R. Gao. From smart grids to an energy internet: Assumptions, architectures and requirements. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, pages 94–98, 2008.

[207] U.S. Department of Energy - Office of Electric Transmission and Distribution. "grid 2030" a national vision for electricity's second 100 years. Technical report, U.S. Department of Energy - Office of Electric Transmission and Distribution, 2003.

[208] V. Vaitheeswaran. *Power to the People*. Earthscan, 2005.

[209] K. H. van Dam and E. J. L. Chappin. Coupling Agent-Based Models of Natural Gas and Electricity Markets. In *First International Workshop on Agent Technologies for Energy Systems, ATES2010, Toronto*, 2010.

[210] A. Vázquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of the Internet. *Physical Review E*, 65(6):1–12, 2002.

[211] T. Verschueren, W. Haerick, K. Mets, C. Develder, F. De Turck, and T. Pollet. Architectures for smart end-user services in the power grid. In *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*, pages 316 –322, 2010.

[212] P. Vytelingum, S. Ramchurn, T. Voice, A. Rogers, and N. Jennings. Agent-based modeling of smart-grid market operations. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–8, 2011.

[213] C. A. Walford. Wind turbine reliability: understanding and minimizing wind turbine operation and maintenance costs. Technical Report SAND2006-1100, Sandia National Laboratries, 2006.

[214] D. Wall, G. Thompson, and J. Northcote-Green. An optimization model for planning radial distribution networks. *Power Apparatus and Systems, IEEE Transactions on*, PAS-98(3):1061 –1068, 1979.

[215] D. Wang, B. de Wit, S. Parkinson, J. Fuller, D. Chassin, C. Crawford, and N. Djilali. A test bed for self-regulating distribution systems: Modeling integrated renewable energy and demand response in the gridlab-d/matlab environment. In *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, pages 1 –7, 2012.

[216] J.-W. Wang and L.-L. Rong. Cascade-based attack vulnerability on the us power grid. *Safety Science*, 47(10):1332–1336, 2009.

[217] Z. Wang, A. Scaglione, and R. Thomas. Electrical centrality measures for electric power grid vulnerability analysis. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 5792 –5797, 2010.

[218] Z. Wang, A. Scaglione, and R. Thomas. Generating statistically correct random topologies for testing smart grid communication and control networks. *Smart Grid, IEEE Transactions on*, 1(1):28–39, 2010.

[219] Z. Wang, A. Scaglione, and R. Thomas. The node degree distribution in power grid and its topology robustness under random and selective node removals. In *Communications Workshops (ICC), 2010 IEEE International Conference on*, pages 1 –5, 2010.

[220] Z. Wang, R. Thomas, and A. Scaglione. Generating random topology power grids. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, page 183, 2008.

[221] C. Warmer, K. Kok, S. Karnouskos, A. Weidlich, and D. Nestle. Web services for integration of smart houses in the smart grid. In *Grid-Interop 2009*, pages 1–5, 2009.

[222] D. J. Watts. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton Univ. Press, Princeton, NJ, USA, 2003.

[223] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[224] D. Whitehead. The el farol bar problem revisited: Reinforcement learning in a potential game. ESE Discussion Papers 186, Edinburgh School of Economics, University of Edinburgh, 2008.

[225] R. Wiser, G. Barbose, and C. Peterman. Tracking the sun: the installed cost of photovoltaics in the u.s. from 1998-2007. Technical report, Lawrence Berkeley National Laboratory, 2009.

[226] Y. Xiaodan, J. Hongjie, W. Chengshan, W. Wei, Z. Yuan, and Z. Jinli. Network reconfiguration for distribution system with micro-grids. In *Sustainable Power Generation and Supply, 2009. SUPERGEN '09. International Conference on*, pages 1 –4, 2009.

[227] Q. Yang, J. Barria, and T. Green. Communication infrastructures for distributed control of power distribution networks. *Industrial Informatics, IEEE Transactions on*, 7(2):316 –327, 2011.

Samenvatting

Het smart grid belooft de manier waarop energie wordt geproduceerd, gedistribueerd en verhandeld te veranderen. Een essentieel punt van het smart grid is het toegevoegde gebruik van ICT-middelen op elk niveau van de energiesector. De toegevoegde informatiestroom die parallel loopt aan de elektrische stroom bevordert de samenwerking tussen de traditionele spelers in de energiesector en de nieuwe rollen die dankzij het smart grid op het energielandschap verschijnen. De eerste reden om het smart grid te implementeren is de vergroting van de flexibiliteit van het energiesysteem, hetgeen nodig is om meer duurzame energiebronnen te kunnen ondersteunen. Een tweede beweegreden voor het smart grid houdt verband met de aanhoudende ontwikkelingen om de energiemarkt uit een monopolie te bevrijden; een proces genaamd *unbundling*. In deze thesis kijken we naar toekomstige scenarios waarin eindgebruikers zowel consumenten alsook, met behulp van kleine duurzame installaties, kleine producenten van energie zijn. Deze nieuwe spelers staan bekend als *prosumenten*. In een dergelijk smart grid scenario onderzoeken we de evolutie van de fysieke infrastructuur vanuit een topologisch oogpunt. Het doel is om de basis van een decision-supportsysteem te verschaffen waarmee beleidsmakers en energiemaatschappijen de kosten en baten van een verandering in het elektriciteitsnetwerk kunnen afwegen. Dit soort nieuwe elektriciteitsnetwerken zullen de gedistribueerde opwekking van energie bevorderen door middel van duurzame bronnen en gelokaliseerde energieverdeling bij de eindgebruikers. Tevens kijken we naar software en applicaties waarmee de wisselwerking tussen de vele spelers op het smart grid gerealiseerd kan worden. Dit is een ander belangrijk punt in het smart grid panorama.

Hoofdstuk 2 vormt een samenvatting van gerelateerd onderzoek. Als eerste is het smart grid concept uitgelegd en zijn de visies van andere wetenschappers in het gebied hierover beschreven. De focus wordt dan verlegd naar de infrastruc-

tuur van elektriciteitsnetwerken, met bijzondere aandacht voor studies die een topologische benadering hebben in het bestuderen van het elektriciteitsnetwerk. De literatuur past een topologische benadering alleen toe op het high voltage netwerk en binnen het traditionele energieparadigma (i.e., geen vermelding van een smart grid). We houden ook rekening met de manier waarop het evolutieproces van een netwerk vanuit een strict graaftheoretisch perspectief en de manier waarop het probleem van elektriciteitsnetwerkuitbreiding en -evolutie worden beschouwd in de literatuur over power engineering. Dit hoofdstuk geeft tevens een overzicht van de applicaties en software-architectuurbenaderingen ten behoeve van het smart grid. We kijken naar de meest veelbelovende architecturen en in het bijzonder naar het service-georiënteerde paradigma. Daarnaast analyseren we marktapplicaties voor het smart grid die in de literatuur zijn voorgesteld.

Hoofdstuk 3 presenteert de analyse van de topologie van midden- en hoogspanningsnetwerken aan de hand van metingen uit Noord-Nederland. Naast een puur topologische analyse hebben we een gewogen analyse uitgevoerd die rekening houdt met de weerstand en capaciteit van kabels (i.e., elektrische stroom). De resultaten onthullen netwerken die topologisch gezien licht verbonden zijn (i.e., low node degree) en minder verbonden zijn dan de hoogspanningsnetwerken. Daarnaast behoren de netwerken niet tot de small-world categorie omdat de clusteringscoefficiënt altijd zeer klein is. In dit hoofdstuk verschaffen we ook een studie waarin verbanden tussen de topologische eigenschappen van het netwerk en de kosten van elektriciteitsdistributie worden gelegd. De eigenschappen die de kosten beïnvloeden kunnen onderverdeeld worden in twee macrocategorien: energieverlies (α metrics) en betrouwbaarheid (β metrics). Onze bevinding is dat de verschillende geanalyseerde monsters verschillend scoren op de metrieken. Daarom zullen in een toekomst waar energie gekocht en verkocht kan worden binnen een wijk of stad, sommige distributienetwerken economisch gezien beter zijn dan anderen.

Het onderzoek naar de evolutie van het elektriciteitsnetwerk binnen een visie die gedistribueerde energieopwekking en gelokaliseerde componenten omvat is beschreven in Hoofdstuk 4. Een set van topologische metrieken wordt geïdentificeerd die de belangrijkste eigenschappen vastleggen waaraan het nieuwe elektriciteitsnetwerk moet voldoen: efficiënte energiedistributie, weerstand tegen fouten en bestendigheid tegen willekeurige en gerichte storingen. Eerst beschouwen we bekende topologieën uit de literatuur over complexe netwerken om het proces waarin netwerkinfrastructuren van de grond af worden opgezet na te bootsen. De uitkomst is dat een kleine verhoging in connectiviteit bevorderlijk is en dat de small-word topologie de beste resultaten geeft volgens de metrieken. Ten tweede kijken we naar de evolutie van huidige delen van het Nederlandse elektriciteitsnetwerk en beschouwen we strategieën om verbindingen toe te voegen waarmee de connectiviteit van de netwerken enigszins kan worden verbeterd. In deze oefeningen modelleren we zowel de fysieke parameters die karakteriserend zijn voor de netwerken alsook de kosten die gemoeid zijn met het bekabelen bij de aanleg van deze infrastructuren. Hieruit volgt dat een goede afweging tussen de topologische prestaties en de uitvoeringskosten gevonden kan worden bij een strategie waarin de dichtstbijzijnde knooppunten met elkaar verbonden worden.

De door ICT ondersteunde aspecten van het smart grid staan beschreven in Hoofdstuk 5. We kijken naar de informatiearchitectuur waarmee op een juiste manier interactie tussen verschillende spelers met verschillende technologieën kan plaatsvinden. We bevinden dat service-georiënteerde architectuur een paradigma is dat goed past bij dit soort heterogene omgevingen. We gaan dieper in op twee applicaties die essentieel zijn voor het toekomstige smart grid: een tool waarmee met de technologieën en services van vandaag, de essentiële mogelijkheden van het smart grid, zoals een dynamische energieprijs, kunnen worden gesimuleerd; ten tweede kijken we naar een applicatie gebaseerd op agenttechnologieën om een markt te creëren waar prosumenten met elkaar kunnen onderhandelen over hun energielevering. Voor de eerstgenoemde applicatie zijn de resultaten van een implementatie die is toegepast in een kantooromgeving toegevoegd. Voor de laatstgenoemde applicatie zijn computersimulaties uitgevoerd.

Het smart grid is een visie die nog moet worden vertaald naar producten en toepassingen. We hebben aspecten van connectiviteit (i.e., topologie) verkend, zowel op infrastructureel als op applicatieniveau, die het mogelijk maken om verder te reizen in de overgang naar het smart grid en om duurzaamheid te vergroten.